

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 28, 2003, 13:27:04 ; Search time 12.5 seconds
(without alignments)

1308.311 Million cell updates/sec

Title: US-10-042-417A-29

Perfect score: 515

Sequence: 1 ccgtactgtgnttcggc.....cgccacacgtgngggcatt 278

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Xgapop 10.0, Xgapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 140709 seqs, 29413474 residues

Total number of hits satisfying chosen parameters: 281418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlh
-Q=/cgn2_1/USPPO.spool/US10042417/runat_19082003_133541_9765/app_query.fasta_1.455
-DB=Pending_Patents_AA_New -QFMT=fastan -SUFFIX=rapn -MINMATCH=0.1 -LOOPCL=0
-LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=DIOSUM62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=ptc -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10042417_CGN_1_1_8 @runat_19082003_133541_9765 -NCPU=6 -ICPU=3
-NO_MMAP -LARGESQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Pending_Patents_AA_New.*

1: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	436.5	84.8	91	6	US-10-042-417A-30
2	355	68.9	404	6	Sequence 3354, Ap
3	349	67.8	403	6	Sequence 1386, Ap
4	229	44.5	44	6	Sequence 63, Appl
5	97	19.9	666	6	US-10-603-113-15507
6	91	18.6	680	7	US-60-487-610-1994
7	89	18.2	317	6	US-10-286-897-5722
8	89	18.2	317	6	US-10-286-897-5723
9	89	18.2	317	6	US-10-288-898A-5722
10	89	18.2	317	6	US-10-258-898A-5723
11	89	18.2	1219	6	US-10-326-956-2641

c 12	88	18.0	424	7	US-60-485-450-1598	Sequence 1598, Ap
c 13	87.5	17.9	515	1	PCT-US02-18638A-156	Sequence 156, App
c 14	87.5	17.9	515	1	PCT-US03-20907-19	Sequence 19, Appl
c 15	87.5	17.9	515	6	US-10-612-090-19	Sequence 19, Appl
c 16	86.5	16.8	297	6	US-10-612-783-4246	Sequence 4246, Ap
c 17	85.5	17.5	491	7	US-60-487-610-1601	Sequence 1601, Ap
c 18	84	17.2	987	7	US-60-487-610-2242	Sequence 2242, Ap
c 19	84	17.2	987	7	US-60-485-450-1437	Sequence 1437, Ap
c 20	83.5	17.1	440	7	US-60-487-610-1602	Sequence 1602, Ap
c 21	83.5	17.1	639	6	US-10-408-765A-2802	Sequence 2802, Ap
c 22	83.5	17.1	984	6	US-10-292-798-936	Sequence 936, App
c 23	82.5	16.9	511	6	US-10-617-320-2904	Sequence 2904, Ap
c 24	82.5	16.9	565	6	US-10-286-897-2346	Sequence 2346, Ap
c 25	82.5	16.9	565	6	US-10-258-898A-2346	Sequence 2346, Ap
c 26	82	16.8	297	6	US-10-612-783-4246	Sequence 4246, Ap
c 27	82	16.8	1318	1	PCT-US03-11231-197	Sequence 197, App
c 28	81.5	15.8	69	6	US-10-286-897-2542	Sequence 2542, Ap
c 29	81.5	15.8	69	6	US-10-258-898A-2542	Sequence 2542, Ap
c 30	81.5	16.7	173	6	US-10-408-765A-69	Sequence 69, Appl
c 31	81.5	15.8	190	6	US-10-286-897-6114	Sequence 6114, Ap
c 32	81.5	15.8	190	6	US-10-258-898A-6114	Sequence 6114, Ap
c 33	81.5	16.7	349	5	US-09-780-996A-7	Sequence 7, Appl
c 34	81.5	16.7	440	7	US-60-490-890-2354	Sequence 2354, Ap
c 35	80.5	16.5	580	6	US-10-273-573-8952	Sequence 8952, Ap
c 36	80.5	16.5	3067	6	US-10-631-467-1618	Sequence 1618, Ap
c 37	80	16.4	148	6	US-10-273-573-5621	Sequence 5621, Ap
c 38	80	15.5	232	6	US-10-292-798-1116	Sequence 1116, Ap
c 39	79	16.2	624	6	US-10-273-573-10422	Sequence 10422, A
c 40	79	16.2	638	6	US-10-372-209-17	Sequence 17, Appl
c 41	79	16.2	702	1	PCT-US03-24669-288	Sequence 288, App
c 42	78.5	16.1	1083	6	US-10-293-244-3567	Sequence 3567, Ap
c 43	78.5	16.1	1151	6	US-10-293-244-1599	Sequence 1599, Ap
c 44	78.5	16.1	1256	1	PCT-US03-18934-111	Sequence 111, App
c 45	78.5	15.2	1564	6	US-10-292-798-1254	Sequence 1254, Ap

ALIGNMENTS

RESULT 1

US-10-042-417A-30

; Sequence 30, Application US/10042417A

; GENERAL INFORMATION:

; APPLICANT: Pagano, M.

; TITLE OF INVENTION: METHODS TO IDENTIFY COMPOUNDS USEFUL FOR THE TREATMENT OF PROLIFERATIVE DISORDERS

; FILE REFERENCE: 5914-090-999

; CURRENT APPLICATION NUMBER: US/10/042.417A

; CURRENT FILING DATE: 2002-01-07

; PRIOR APPLICATION NUMBER: 60/260,179

; PRIOR FILING DATE: 2001-01-05

; NUMBER OF SEQ ID NOS: 92

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 30

; LENGTH: 91

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: SITE

; LOCATION: all xaa positions

; OTHER INFORMATION: Xaa=unknown amino acid residue

US-10-042-417A-30

Alignment Scores:
Pred No.: 1-42e-33 Length: 91
Score: 436.50 Matches: 91
Percent Similarity: 98.91% Conservative: 0
Best Local Similarity: 98.91% Mismatches: 0
Query Match: 84.76% Indels: 1
DB: 6 Gaps: 1

US-10-042-417A-29 (1-278) x US-10-042-417A-30 (1-91)

Qy 2 CGTAGTACTGGNTTCCGGCGGCTGGTGGAGGAATGGAGCCGTAGTGTTCGGCGAGT 61

Db		1	ArgSerThr.GlyPheArgAlaGluIrrPserArg----	LeuAlaAlaSer	19
Qy		62	CCCGGNTCCCTCGTAGCCCGCGGANNACCTTCGTGTGAGTAACCTGCGGAGGTGGTG	121	
Db		20	ProGly***LeuArgArgProAla***ThrPheValLeuSerAsnLeuAlaGluValVal	39	
Qy		122	GAGCGTGTCTCACCTTCTCTGCCGCCAAGCGTGTGTCGGGTGGCTCGCGTGTGCCGC	181	
Db		40	GluArgValLeuThrPheLeuProAlaLysAlaLeuLeuArgValAlaCysValCysArg	59	
Qy		182	TTATGAGGAGGAGTGTGCGCAGAGTATTGCGGACCCTCGGAGCGTAACCTGGATCTCC	241	
Db		60	LeuTrpArgGluCysValArgValLeuArgThrHisArgSerValThrTrpIleSer	79	
Qy		242	GCAGGCCTGGCGGAGCCGCCACCTGGNGGGGCAT	277	
Db		80	AlaGlyLeuAlaGluAlaGlyHisLeu***GlyHis	91	

```

RESULT 2
US-10-293-244-3354
: Sequence 3354, Application US/10293244
: GENERAL INFORMATION:
: APPLICANT: Hyseq, Inc.
: APPLICANT: Tang, Y. Tom et al
: TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
: FILE REFERENCE: 21272-029
: CURRENT APPLICATION NUMBER: US/10/293,244

```

Alignment Scores:		
Pred. No.:	6.5e-26	404
Score:	355.00	77
Percent Similarity:	92.77%	Conservative: 0
Best Local Similarity:	92.77%	Mismatches: 6
Query Match:	68.93%	Indels: 2
DB:	6	Gaps: 0

US-10-042-417A-29 (1-278) x US-10-293-244-3354 (1-404)

31	QY	GGATGGAGCCGGTAGNTGCTTGGCGGAGATCCCG - GNTCTCCGTAGACCCGCGGAN -	88
1	Db	GGATGGAGCCGGTAGNTGCTTGGCGGAGATCCCG - GNTCTCCGTAGACCCGCGGAN -	88
1	Db	GlyMetGluProValGlyCysGlyCysGlySerValAspProArgSer 20	
89	QY	ACCTTCGTGTGAGTAACCTGGCGGAGGTGTGGAGCGTGTGTACCTTCCTGCCCGCC	148
21	Db	ThrPheValLeuSerAsnLeuAlaGluValValGluArgValLeuThrPheLeuProAla 40	

Qy	149	AAGCGTTCCTCGCGGTCGCGCTGTCGCGCTTATCGAGGGAGTGCTGTGCGCAGAGTA	208
Db	41	LysAlaLeuLeuArgValAlaCysValCysArgLeuTrpArgGluCysValArgArgVal	60
Qy	209	TTGCGGACCATCGAGGCGTAACCTGGATCTCCGACGCGCTGGCGGAGCGCGCCACCTG	268
Db	61	LeuArgThrHisArgSerValThrTrpIleSerAlaGlyLeuAlaGluAlaGlyHisLeu	80
Qy	269	GNGGGGCAT	277
Db	81	GluGluHis	83

```

RESULT 3
US-10-293-244-1386
; Sequence 1386, Application US/10293244
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang, Y. Tom et al
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-029
; CURRENT APPLICATION NUMBER: US/10/293,244
; CURRENT FILING DATE: 2002-11-12

```

Alignment Scores:		
Pred. No.:	2,378-25	Length:
Score:	349.00	Matches:
Percent Similarity:	92.68%	Conservative:
Best Local Similarity:	92.68%	Mismatches:
Query Match:	67.77%	Indels:
DB:	6	Gaps:
		403

US-10-042-417A-29 (1-278) x US-10-293-244-1386 (1-403)

34	ATGGAGCCGGTAGNTGCTTGGCGGAGATCCCG - GGNTCCTCCGAGACCCGCCGAN - ACC	91
QY		
1	MetGluProValGlyCysCysGlyGluCysArgGlySerValAspProArgSerThr	20
Db		
92	TTCTGTTTGAGTAACCTGGCGGAGGTGGTGGAGCGGTGCTCACCTTCCTGCCCGCCCAAG	151
QY		
21	PheValLeuSerAsnLeuAlaGluValValGluArgValLeuThrPheLeuProAlaLys	40
Db		
152	CGCTTGCTGCGGGTGGCTGCTGCGCGCTTATGGAGGAGTGTGTGCCACAGATATTG	211
QY		
41	AlaLeuLeuArgValAlaCysValCysArgLeuTrpargGluCysValArgArgValLeu	60
Db		
212	CGGACCCATCGAGCGTAACCTGGATCTCCGCAAGGCTGCGGAGCGGCCACCTGGNG	271
QY		
61	ArgThrHisArgSerValThrTrpIleSerAlaGlyLeuAlaGluAlaGluHisLeuGlu	80
Db		


```
QY 272 GGCAT 277
Db 81 GlyHis 82

-RESULT 4
US-10-042-417A-63
; Sequence 63, Application US/10042417A
; GENERAL INFORMATION:
; APPLICANT: Pagano, M.
; TITLE OF INVENTION: METHODS TO IDENTIFY COMPOUNDS USEFUL FOR THE TREATMENT OF PROLIFERATIVE DISORDERS
; FILE REFERENCE: 5914-090-999
; CURRENT APPLICATION NUMBER: US/10/042.417A
; CURRENT FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: 60/260,179
; PRIOR FILING DATE: 2001-01-05
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 63
; LENGTH: 44
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-042-417A-63

Alignment Scores:
Pred. No.: 3,76e-14 Length: 44
Score: 229.00 Matches: 44
Percent Similarity: 100.00% Conservative: 44
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 44.47% Indels: 0
DB: 6 Gaps: 0

US-10-042-417A-29 (1-278) x US-10-042-417A-63 (1-44)
QY 107 CTGCGGAGTGGTGGAGCGTGTCTCACCTTCTGCGCCGCAAGGCGTGTCTGCGGGTG 166
Db 1 LeuAlaGluValValGluArgValLeuThrPheLeuProAlaLysAlaLeuLeuArgVal 20
QY 167 GCCTGCGTGTGCCCTTATGAGGAGTGTGTGCGCAGAGTATTGCGGACCATCGGAGC 226
Db 21 AlacysValCysArgLeuTrpArgGluCysValargValLeuArgThrHisArgSer 40
QY 227 GTAACCTGGATC 238
Db 41 ValThrTrpIle 44

-RESULT 5
US-10-603-113-15507
; Sequence 15507, Application US/10603113
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/10/603.113
; CURRENT FILING DATE: 2003-06-24
; PRIOR APPLICATION NUMBER: US/09/248,796
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 28206
; SEQ ID NO 15507
; LENGTH: 666
; TYPE: PRT
; ORGANISM: Candida albicans
US-10-603-113-15507

Alignment Scores:
Pred. No.: 0.0975 Length: 666
Score: 97.00 Matches: 33
Percent Similarity: 42.11% Conservative: 7
Best Local Similarity: 34.74% Mismatches: 47
Query Match: 19.88% Indels: 8
DB: 6 Gaps: 3

US-10-042-417A-29 (1-278) x US-10-603-113-15507 (1-666)
QY 272 CCNCCAGTGGCGCGG-----CCTCGCCAGCGCTCGGAGATCCAGGTGA 228
Db 487 ProProAlaAlaArgAlaThrProProAlaProProAlaArgAlaThrAsnGlnLeu 506
QY 227 CGCTCCGATGGTCCGCAATATCTTCGCACACACTCCATCAATAGCGGCACGACG 168
Db 507 GlyLeuProProArgAsnThrLeu-----GlyLeuProProArgThrAsnAsnThrGln 524
QY 167 CCACCCGGAGCAAGCGCTTGGCGGAGGAGGTGAGCACAGCTCCACCTCCGCGCA 108
Db 525 ProProAlaProProArgAlaSerArgGlyAlaValProProProProPro 544
QY 107 GGTACTTCAACACGAAGGTNTCCGCGGTCTACGGAGGAGGAGGAGTCCGCGCAACA 48
Db 545 ProProAlaThrArgAlaPro---MetGlnLeuGlnLeuGlnSerSerProGlnIns 563
QY 47 NCTACCGGTCTTCCTCACCAGCGCCGCGGAAANCCAGTACTAC 3
Db 564 SerProIleSerProAlaGlnGlnGlnGlnGlnTyr 578

-RESULT 6
US-60-487-610-1994
; Sequence 1994, Application US/60487610
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: HUANG, Hongjin
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INFECTED SUBJECTS,
; TITLE OF INVENTION: METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001469
; CURRENT APPLICATION NUMBER: US/60/487,610
; CURRENT FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 97101
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1994
; LENGTH: 680
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-487-610-1994

Alignment Scores:
Pred. No.: 0.356 Length: 680
Score: 91.00 Matches: 34
Percent Similarity: 40.00% Conservative: 6
Best Local Similarity: 34.00% Mismatches: 39
Query Match: 18.65% Indels: 21
DB: 7 Gaps: 4

US-10-042-417A-29 (1-278) x US-60-487-610-1994 (1-680)
QY 272 CCNCCAGTGGCGCGCTTCGCCAGCGCTCGGAGATCCAGGTACGCTCCGATGGGTCC 213
Db 248 ProProGlyProGlnGlyProGlyGluArgGlyProGlyGlyIleGlyLysProGly 267
QY 212 GCAATACTCTCGGCACACACTCCCTCCATAGCGGCACACGAGCGCCACGCGACAG 153
Db 268 AlaAlaGlyAlaProGlyGlnProGlyIleProGlyThrLysGlyLeuProGlyAla 286
QY 152 CCTTGGCGGAGGAAGGTGAGCACGCTCCACACCTCCGCGAGTTACTCAACAGA 93
Db 287 ProGlyIleAlaGly-----ProProGlyProProGlyPheGlyLysPro 301
QY 92 AGGTNTCCGCGGTCTAC-----GGAGGAGNCCGGGAC 60
Db 302 GlyLeuProGlyLeuLysGlyGluArgGlyProAlaGlyLeuProGlyGly-ProGlyAl 321
QY 59 TCGCCGCAAGCANTACCGGCT-----CCATTCTCCACGCGCGCG 17
Db 321 aLysGlyGluGlnGlyProAlaGlyLeuProGlyLysProGlyLeuThrGlyProPro 340
```



```
QY 119 CCACCTCGCCCA 108
    |||||
Db 1074 ProProProPro 1077

RESULT 12
US-60-485-450-1598
; Sequence 1598, Application US/60485450
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: CHANG, Sheng-Yung
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RESPONSE TO INTERFERON TREATMENT IN HEPATITIS C
; TITLE OF INVENTION: VIRUS-INFECTED SUBJECTS, METHODS OF DETECTION AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001470
; CURRENT APPLICATION NUMBER: US/60/485,450
; CURRENT FILING DATE: 2003-07-09
; NUMBER OF SEQ ID NOS: 47859
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1598
; LENGTH: 424
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-485-450-1598

Alignment Scores:
Pred. No.: 0.666 Length: 424
Score: 88.00 Matches: 29
Percent Similarity: 41.67% Conservatives: 6
Best Local Similarity: 34.52% Mismatches: 35
Query Match: 18.03% Indels: 14
DB: 7 Gaps: 5

US-10-042-417A-29 (1-278) x US-60-485-450-1598 (1-424)
QY 272 CCNCCAGTGGCGCGCTCCGCCAGGCTGCGGAGATCCAGGTACGCTCGGATGGTCC 213
    |||||
Db 297 ProProGlyGlyMetProHisProGlyMetSerGlnMetGlnLeuAlaHisGlyPro 316
    |||||

QY 212 GCAATACCTCTCGGCACACACT-----CCCTCCATAAGCGGCACACGCGCCACCC 162
    |||||
Db 317 HisGlyLeu---GlyHisProHisAlaGlyProGlySerGlyGlyGlnProProPro 335
    |||||

QY 161 GCACGACGCTTGGCGGGCAGGAAGGTGAGCAGCCTCCACACCTCGGCCAGGTAC 102
    |||||
Db 336 ArgProProGlyMet-----ProHisProGlyProProProMetGlyMet 351
    |||||

QY 101 TCACACGAGGTTCCGCGGGTCTACGGAGGAGNCCCGGACTCGCGCGCAGCAGCTACC 42
    |||||
Db 352 ProProArgGlyProPro-----PheGly-----SerProMetGlyHisPro 365
    |||||

QY 41 GGCCTCCATTCTCT 30
    |||||
Db 366 GlyProMetPro 369

RESULT 13
PCT-US02-18638A-156
; Sequence 156, Application PC/TUS0218638A
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc. et al.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF CERVICAL CANCER
; FILE REFERENCE: MRI-035PC
; CURRENT APPLICATION NUMBER: PCT/US02/18638A
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14

; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 156
; LENGTH: 515
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-18638A-156

Alignment Scores:
Pred. No.: 0.748 Length: 515
Score: 87.50 Matches: 32
Percent Similarity: 35.29% Conservatives: 4
Best Local Similarity: 31.37% Mismatches: 30
Query Match: 17.93% Indels: 36
DB: 1 Gaps: 5

US-10-042-417A-29 (1-278) x PCT-US02-18638A-156 (1-515)
QY 269 CCAGGTGGCGCGCTCCGCCAGGCTGCGGAGATCCAGGTACGCTCGGATGGTCCGCA 210
    |||||
Db 122 ProAspAsnLysProAlaPro-----GlySerThr 131
    |||||

QY 209 ATACTCTGCGCACAC---ACTCCCTCCATAAGCGGCACACGCGCCACCGCCGACGACG 153
    |||||
Db 132 AlaProProAlaHisGlyValThrSerAlaProAspThrArgProProGlySerThr 151
    |||||

QY 152 CTTTGGCGGCGGAGGTGAGCAGCAGCTCCACCA----- 117
    |||||
Db 152 -----AlaProProAlaHisGlyValThrSerAlaPro 162
    |||||

QY 116 -----CCTCGCGCAGGTACTCAACACAGGAAGTNTCCGCGGGTCTACGGAGGANC 66
    |||||
Db 163 AspThrArgProProProGlySerThrAla-ProAlaAlaHisGlyValThrSerAlaPr 182
    |||||

QY 65 CGGAGCTCGCGGCGGAGCAGTACCGGCTCCATTCTCACCAGCCGCGCGGAGNCCAGTAC 6
    |||||
Db 182 oAspThrArg-----ProAlaProGlySerThrAlaProProAlaHisGlyVal 198
    |||||

QY 5 TACG 2
    |||||
Db 198 lThr 199

RESULT 14
PCT-US03-20907-19
; Sequence 19, Application PC/TUS0320907
; GENERAL INFORMATION:
; APPLICANT: Immunogen, Inc.
; TITLE OF INVENTION: ANTIBODIES TO NON-SHED MUC1 AND MUC16, AND USES THEREOF
; FILE REFERENCE: F167822
; CURRENT APPLICATION NUMBER: PCT/US03/20907
; CURRENT FILING DATE: 2003-07-03
; PRIOR APPLICATION NUMBER: US 60/393,094
; PRIOR FILING DATE: 2002-07-03
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 19
; LENGTH: 515
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Exemplary Muc1 protein
PCT-US03-20907-19

Alignment Scores:
Pred. No.: 0.748 Length: 515
Score: 87.50 Matches: 32
Percent Similarity: 35.29% Conservatives: 4
Best Local Similarity: 31.37% Mismatches: 30
Query Match: 17.93% Indels: 36
DB: 1 Gaps: 5

US-10-042-417A-29 (1-278) x PCT-US03-20907-19 (1-515)
```

QY 269 CCAGGTGGCGGCTCCGCCAGGCTCGGAGATCCAGGTACCTCCGATGGGTCCGCA 210
Db ||| ::||| ||| |||||
122 ProAspAsnLysProAlaPro-----GlySerThr 131
QY 209 ATACTCTGGCGACAC---ACTCCCTCCATAGCGGCACACGCGGCCACCGCAGCAACG 153
Db ||||| ||| ||| ||||| ||||| ::|||
132 AlaProProAlaHisGlyValThrSerAlaProAspThrArgProProGlySerThr 151
QY 152 CTTGGCGGCGAGGAGTGTCAGCACAGCTCCACCA----- 117
Db ||||| |||||
152 -----AlaProProAlaHisGlyValThrSerAlaPro 162
QY 116 -----CCTCGCGCAGGTTACTCAACACAGAGAGTNTCCGGGGTCTACGGAGGANC 66
Db ||||| ||||| ::||| ||| ||| |||
163 AspThrArgProProGlySerThrAla-ProAlaAlaHisGlyValThrSerAlaPr 182
QY 65 CGGACTCGCGCGAAGCAGTACCGCTCCATTCCTCACCAGCCCGCGGAANCCAGTAC 6
Db ||||| ||||| ||||| |||::|||
182 oAspThrArg-----ProAlaProGlySerThrAlaProProAlaHisGlyVa 198
QY 5 TACG 2
Db |||
198 lThr 199

RESULT 15

US-10-612-090-19
; Sequence 19, Application US/10612090
; GENERAL INFORMATION:
; APPLICANT: Immunogen, Inc.
; TITLE OF INVENTION: ANTIBODIES TO NON-SHED MUC1 AND MUC16, AND USES THEREOF
; FILE REFERENCE: A8340
; CURRENT APPLICATION NUMBER: US/10/612,090
; CURRENT FILING DATE: 2003-07-03
; PRIOR APPLICATION NUMBER: US 60/393,094
; PRIOR FILING DATE: 2002-07-03
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 19
; LENGTH: 515
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Exemplary Muc1 protein
US-10-612-090-19

Alignment Scores:
Pred. No.: 0.748 Length: 515
Score: 87.50 Matches: 32
Percent Similarity: 35.29% Conservative: 4
Best Local Similarity: 31.37% Mismatches: 30
Query Match: 17.93% Indels: 36
DB: 6 Gaps: 5

US-10-042-417A-29 (1-278) x US-10-612-090-19 (1-515)

QY 269 CCAGGTGGCGGCTCCGCCAGGCTCGGAGATCCAGGTACCTCCGATGGGTCCGCA 210
Db ||| ::||| ||| |||||
122 ProAspAsnLysProAlaPro-----GlySerThr 131
QY 209 ATACTCTGGCGACAC---ACTCCCTCCATAGCGGCACACGCGGCCACCGCAGCAACG 153
Db ||||| ||| ||| ||||| ||||| ::|||
132 AlaProProAlaHisGlyValThrSerAlaProAspThrArgProProGlySerThr 151
QY 152 CTTGGCGGCGAGGAGTGTCAGCACAGCTCCACCA----- 117
Db ||||| |||||
152 -----AlaProProAlaHisGlyValThrSerAlaPro 162
QY 116 -----CCTCGCGCAGGTTACTCAACACAGAGTNTCCGGGGTCTACGGAGGANC 66
Db ||||| ||||| ::||| ||| ||| |||
163 AspThrArgProProGlySerThrAla-ProAlaAlaHisGlyValThrSerAlaPr 182
QY 65 CGGACTCGCGCGAAGCAGTACCGCTCCATTCCTCACCAGCCCGCGGAANCCAGTAC 6
Db ||||| ||||| ||||| |||::|||

Db 182 oAspThrArg-----ProAlaProGlySerThrAlaProProAlaHisGlyVa 198
QY 5 TACG 2
Db |||
198 lThr 199
Search completed: August 28, 2003, 13:37:57
Job time : 17.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 28, 2003, 13:25:25 ; Search time 233.5 Seconds
(without alignments)
2072.458 Million cell updates/sec

Title: US-10-042-417A-29

Sequence: 1 cctagtactgnttcggc.....cggccacctgnggggcatt 278

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5580241 seqs, 870357830 residues

Total number of hits satisfying chosen parameters: 11160482

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlh
-Q/cgn2_1/USPRO.spool/US10042417/runat_19082003_133540_9700/app.query.fasta.1.455
-DB=Pending_Patents_AA_Main -QMT=fastan -SUFFIX=rapm -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blom62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10042417 -CGN_1_1_37 -runat_19082003_133540_9700 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DSLOP=6 -DELEXT=7

Database :

Pending_Patents_AA_Main:*
1: /cgn2_6/ptodata/2/paa/PCTUS_COMB.pcp.*
2: /cgn2_6/ptodata/2/paa/US06_COMB.pcp.*
3: /cgn2_6/ptodata/2/paa/US07_COMB.pcp.*
4: /cgn2_6/ptodata/2/paa/US08_COMB.pcp.*
5: /cgn2_6/ptodata/2/paa/US081_COMB.pcp.*
6: /cgn2_6/ptodata/2/paa/US082_COMB.pcp.*
7: /cgn2_6/ptodata/2/paa/US083_COMB.pcp.*
8: /cgn2_6/ptodata/2/paa/US084_COMB.pcp.*
9: /cgn2_6/ptodata/2/paa/US085_COMB.pcp.*
10: /cgn2_6/ptodata/2/paa/US086_COMB.pcp.*
11: /cgn2_6/ptodata/2/paa/US087_COMB.pcp.*
12: /cgn2_6/ptodata/2/paa/US088_COMB.pcp.*
13: /cgn2_6/ptodata/2/paa/US089_COMB.pcp.*
14: /cgn2_6/ptodata/2/paa/US090_COMB.pcp.*
15: /cgn2_6/ptodata/2/paa/US091_COMB.pcp.*
16: /cgn2_6/ptodata/2/paa/US092_COMB.pcp.*
17: /cgn2_6/ptodata/2/paa/US093_COMB.pcp.*
18: /cgn2_6/ptodata/2/paa/US094_COMB.pcp.*
19: /cgn2_6/ptodata/2/paa/US095_COMB.pcp.*
20: /cgn2_6/ptodata/2/paa/US096_COMB.pcp.*
21: /cgn2_6/ptodata/2/paa/US097A_COMB.pcp.*
22: /cgn2_6/ptodata/2/paa/US097B_COMB.pcp.*
23: /cgn2_6/ptodata/2/paa/US098_COMB.pcp.*
24: /cgn2_6/ptodata/2/paa/US099A_COMB.pcp.*
25: /cgn2_6/ptodata/2/paa/US099B_COMB.pcp.*
26: /cgn2_6/ptodata/2/paa/US100_COMB.pcp.*
27: /cgn2_6/ptodata/2/paa/US101_COMB.pcp.*
28: /cgn2_6/ptodata/2/paa/US102_COMB.pcp.*

29: /cgn2_6/ptodata/2/paa/US103_COMB.pcp.*
30: /cgn2_6/ptodata/2/paa/US104_COMB.pcp.*
31: /cgn2_6/ptodata/2/paa/US60_COMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	436.5	84.8	91	1	PCT-US02-00311-30
2	436.5	84.8	91	1	PCT-US99-19560-30
3	436.5	84.8	91	17	US-09-385-219-30
4	436.5	84.8	91	17	US-09-385-219A-30
5	436.5	84.8	91	26	US-10-042-417-30
6	355	68.9	404	1	PCT-US01-03800A-2042
7	355	68.9	404	1	PCT-US01-04098A-3354
8	355	68.9	404	28	US-10-276-774-2042
9	349	67.8	276	20	US-09-629-469A-16074
10	349	67.8	276	27	US-10-153-668-162
11	349	67.8	277	28	US-10-219-051B-5836
12	349	67.8	277	28	US-10-219-051B-5839
13	349	67.8	277	28	US-10-219-051B-11414
14	349	67.8	403	1	PCT-US01-04098A-1386
15	349	67.8	403	1	PCT-US02-11152-15
16	349	67.8	403	27	US-10-153-668-160
17	349	67.8	403	28	US-10-287-218-15
18	229	44.5	44	1	PCT-US02-00311-63
19	229	44.5	44	1	PCT-US99-19560-63
20	229	44.5	44	17	US-09-385-219-63
21	229	44.5	44	17	US-09-385-219A-63
22	229	44.5	44	26	US-10-042-417-63
23	108.5	22.2	369	30	US-10-437-963-196554
24	104.5	21.4	125	21	US-09-708-427-57794
25	102.5	19.9	192	22	US-09-758-472-7390
26	102.5	19.9	192	28	US-10-235-926-7390
27	102	20.9	794	27	US-10-179-131-6755
28	101.5	20.8	353	30	US-10-437-963-320523
29	101	20.7	162	30	US-10-419-128-32101
30	100.5	19.5	328	30	US-10-419-128-26586
31	100.5	20.6	472	30	US-10-437-963-136000
32	99.5	20.4	98	21	US-09-708-427-84830
33	99.5	20.4	238	30	US-10-437-963-138837
34	99	20.3	255	30	US-10-425-114-47883
35	98.5	20.2	185	30	US-10-437-963-141398
36	98.5	20.2	1269	30	US-10-437-963-190985
37	98	19.0	161	1	PCT-US01-08631-34478
38	98	19.0	161	1	PCT-US01-08631-34481
39	98	20.1	222	21	US-09-708-427-55203
40	98	20.1	409	30	US-10-437-963-140753
41	97.5	20.0	167	21	US-09-733-089-21881
42	97.5	20.0	167	23	US-09-816-660-21881
43	97.5	20.0	258	30	US-10-419-128-16884
44	97	19.9	408	30	US-10-437-963-107919
45	97	19.9	533	30	US-10-437-963-176617

ALIGNMENTS

RESULT 1

PCT-US02-00311-30
; Sequence 30, Application PC/TUS0200311
; GENERAL INFORMATION:
; APPLICANT: Pagano, N.
; TITLE OF INVENTION: METHODS TO IDENTIFY COMPOUNDS USEFUL FOR THE TREATMENT OF PROLIFERATIVE AND DIFFERENTIATIVE DISORDERS
; FILE REFERENCE: 5914-090-228
; CURRENT APPLICATION NUMBER: PCT/US02/00311
; CURRENT FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: 60/260,179

;; PRIOR FILING DATE: 2001-01-5
;; NUMBER OF SEQ ID NOS: 89
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 30
;; LENGTH: 91

;; TYPE: PRT
;; ORGANISM: Homo sapiens

;; FEATURE:

;; NAME/KEY: SITE

;; LOCATION: all Xaa positions

;; OTHER INFORMATION: Xaa=unknown amino acid residue

PCT-US02-00311-30

Alignment Scores:
Pred. No.: 4.25e-30 Length: 91
Score: 436.50 Matches: 91
Percent Similarity: 98.91% Conservatives: 0
Best Local Similarity: 98.91% Mismatches: 0
Query Match: 84.76% Indels: 1
DB: 1 Gaps: 1

US-10-042-417A-29 (1-278) x PCT-US02-00311-30 (1-91)

Qy 2 CGTAGTACTGNTTCCGGCGGCTGGTGAGGAATGGAGCGGTAGTGTCTTGGCGGAGT 61
Db 1 ArgSerThrGlyPheArgArgAlaGlyGluGluTrpSerArg-----LeuAlaAlaSer 19
Qy 62 CCCGGNTCCCTCCGTAGACCCCGGAGANACCTTCGTGTGAGTAACCTGGCGGAGGTGGTG 121
Db 20 ProGly***LeuArgArgProAla***ThrPheValLeuSerAsnLeuAlaGluVal 39
Qy 122 GAGCGTGTCTACCTTCCTCCCGCCGCAAGCGGTGTGCTGGGGTGGCTGGCTGTGCCGC 181
Db 40 GluArgValLeuThrPheLeuProAlaLysAlaLeuArgValAlaCysValCysArg 59
Qy 182 TTATGGAGGGAGTGTGTCCGAGAGTATTGGGACCCATCGGACCGCTAACTGGATCTCC 241
Db 60 LeuTrpArgGluCysValArgArgValLeuArgThrHisArgSerValThrTrpIleSer 79
Qy 242 GCAGCGCTGGCGGAGCGCGCCACTGGNGGGCAT 277
Db 80 AlaGlyLeuAlaGluAlaGlyHisLeu***GlyHis 91

RESULT 2

PCT-US99-19560-30

;; Sequence 30, Application PC/TUS9919560

;; GENERAL INFORMATION:

;; APPLICANT: NEW YORK UNIVERSITY

;; TITLE OF INVENTION: NOVEL UBIQUITIN LIGASES AS THERAPEUTIC TARGETS

;; FILE REFERENCE: 5914-081-228

;; CURRENT APPLICATION NUMBER: PCT/US99/19560

;; CURRENT FILING DATE: 1999-08-31

;; EARLIER APPLICATION NUMBER: 60/098,355

;; EARLIER FILING DATE: 1998-08-28

;; EARLIER APPLICATION NUMBER: 60/118,568

;; EARLIER FILING DATE: 1999-02-03

;; EARLIER APPLICATION NUMBER: 60/124,449

;; EARLIER FILING DATE: 1999-03-15

;; NUMBER OF SEQ ID NOS: 89

;; SOFTWARE: PatentIn Ver. 2.0

;; SEQ ID NO 30

;; LENGTH: 91

;; TYPE: PRT

;; ORGANISM: Homo sapiens

;; FEATURE:

;; NAME/KEY: SITE

;; LOCATION: all xaa positions

;; OTHER INFORMATION: Xaa=unknown amino acid residue

PCT-US99-19560-30

Alignment Scores:
Pred. No.: 4.25e-30 Length: 91
Score: 436.50 Matches: 91

Percent Similarity: 98.91% Conservatives: 0
Best Local Similarity: 98.91% Mismatches: 0
Query Match: 84.76% Indels: 1
DB: 1 Gaps: 1

US-10-042-417A-29 (1-278) x PCT-US99-19560-30 (1-91)

Qy 2 CGTAGTACTGNTTCCGGCGGCTGGTGAGGAATGGAGCGGTAGTGTCTTGGCGGAGT 61
Db 1 ArgSerThrGlyPheArgArgAlaGlyGluGluTrpSerArg-----LeuAlaAlaSer 19
Qy 62 CCCGGNTCCCTCCGTAGACCCCGGAGANACCTTCGTGTGAGTAACCTGGCGGAGGTGGTG 121
Db 20 ProGly***LeuArgArgProAla***ThrPheValLeuSerAsnLeuAlaGluVal 39
Qy 122 GAGCGTGTCTACCTTCCTCCCGCCGCAAGCGGTGTGCTGGGGTGGCTGGCTGTGCCGC 181
Db 40 GluArgValLeuThrPheLeuProAlaLysAlaLeuArgValAlaCysValCysArg 59
Qy 182 TTATGGAGGGAGTGTGTCCGAGAGTATTGGGACCCATCGGACCGCTAACTGGATCTCC 241
Db 60 LeuTrpArgGluCysValArgArgValLeuArgThrHisArgSerValThrTrpIleSer 79
Qy 242 GCAGCGCTGGCGGAGCGCGCCACTGGNGGGCAT 277
Db 80 AlaGlyLeuAlaGluAlaGlyHisLeu***GlyHis 91

RESULT 3

US-09-385-219-30

;; Sequence 30, Application US/09385219

;; GENERAL INFORMATION:

;; APPLICANT: Chitaur, D.

;; APPLICANT: Pagano, M.

;; APPLICANT: Latres, E.

;; TITLE OF INVENTION: NOVEL UBIQUITIN LIGASES AS THERAPEUTIC TARGETS

;; CURRENT APPLICATION NUMBER: US/09/385,219

;; CURRENT FILING DATE: 1999-08-27

;; EARLIER APPLICATION NUMBER: 60/098,355

;; EARLIER FILING DATE: 1998-08-28

;; EARLIER APPLICATION NUMBER: 60/118,568

;; EARLIER FILING DATE: 1999-02-03

;; EARLIER APPLICATION NUMBER: 60/124,449

;; EARLIER FILING DATE: 1999-03-15

;; NUMBER OF SEQ ID NOS: 89

;; SOFTWARE: PatentIn Ver. 2.0

;; SEQ ID NO 30

;; LENGTH: 91

;; TYPE: PRT

;; ORGANISM: Homo sapiens

;; FEATURE:

;; NAME/KEY: SITE

;; LOCATION: all xaa positions

;; OTHER INFORMATION: Xaa=unknown amino acid residue

US-09-385-219-30

Alignment Scores:

Pred. No.: 4.25e-30 Length: 91

Score: 436.50 Matches: 91

Percent Similarity: 98.91% Conservatives: 0

Best Local Similarity: 98.91% Mismatches: 0

Query Match: 84.76% Indels: 1

DB: 1 Gaps: 1

US-10-042-417A-29 (1-278) x US-09-385-219-30 (1-91)

Qy 2 CGTAGTACTGNTTCCGGCGGCTGGTGAGGAATGGAGCGGTAGTGTCTTGGCGGAGT 61
Db 1 ArgSerThrGlyPheArgArgAlaGlyGluGluTrpSerArg-----LeuAlaAlaSer 19
Qy 62 CCCGGNTCCCTCCGTAGACCCCGGAGANACCTTCGTGTGAGTAACCTGGCGGAGGTGGTG 121
Db 20 ProGly***LeuArgArgProAla***ThrPheValLeuSerAsnLeuAlaGluVal 39

RESULT 5

```

RESULT-6
PCT-US01-03800A-2042
; Sequence 2042, Application PC/TUS0103800A
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang, Y. Tom et al
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-030
; CURRENT APPLICATION NUMBER: PCT/US01/03800A
; CURRENT FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 2700
; SOFTWARE: Custom
; SEQ ID NO 2042
; LENGTH: 404
; TYPE: PRT
; ORGANISM: Homo sapiens

```



```
Db      81 GluGlyHis 83
|||||
RESULT 9
US-09-629-469A-16074
; Sequence 16074, Application US/09629469A
; GENERAL INFORMATION:
; APPLICANT: OTA, TOSHIO
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: NISHIKAWA, TETSUO
; APPLICANT: HAYASHI, KOJI
; APPLICANT: SAITO, KAORU
; APPLICANT: YAMAMOTO, JUNICHI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: WAKAMATSU, AI
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: OTSUKI, TETSUJI
; TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH CDNA AND THEIR USE
; FILE REFERENCE: 084335/0123
; CURRENT APPLICATION NUMBER: US/09/629,469A
; CURRENT FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: JP 1999-248036
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: JP 1999-300253
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: JP 2000-118776
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183767
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: JP 2000-241899
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/159,590
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 60/183,322
; PRIOR FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 19025
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16074
; LENGTH: 276
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-629-469A-16074

Alignment Scores:
Pred. No.:      2,52e-22      Length:      276
Score:          349.00        Matches:      76
Percent Similarity: 92.68%    Conservative: 0
Best Local Similarity: 92.68% Mismatches:      6
Query Match:    67.77%       Indels:       2
DB:             20           Gaps:         0

US-10-042-417A-29 (1-278) x US-09-629-469A-16074 (1-276)
QY      34 ATGAGCGCGTAGTGTTCGGCGAGTCCCG-GGNTCTCCGTAGACCGCGGAN-ACC 91
Db      1 MetGluProValGlyCysGlyGluCysArgGlySerValAspProArgSerThr 20
QY      92 TTCGTGTTAGTAACCTGCGGAGGTGGTGGAGCGTGTCTACCTTCCTGCCGCCCAAG 151
Db      21 PheValLeuSerAsnLeuAlaGluValGluArgValLeuThrPheLeuProAlaLys 40
QY      152 GCGTTGCTCGGGTGGCTGCTGCTGCTATGGAGGGAGTGTGTCGCAGAGTATTG 211
Db      41 AlalaLeuArgValAlaCysValCysArgLeuTrpArgGluCysValArgValLeu 60
QY      212 CGGACCCATCGAGCGTAACCTGCATCTCCGAGCGCTGGCGAGCGCGCCACCTGGNG 271
Db      61 ArgThrHisArgSerValThrTrpIleSerAlaGlyLeuAlaGlyHisLeuGlu 80
QY      272 GGGCAT 277
Db      81 GlyHis 82

RESULT 11
US-10-219-051B-5836
; Sequence 5836, Application US/10219051B
; GENERAL INFORMATION:
; APPLICANT: The General Hospital Corporation doing business as Massachusetts General
; APPLICANT: Hospital / Bayer AG
; TITLE OF INVENTION: Nucleotide sequences involved in pain
; FILE REFERENCE: Lea 35693 Foreign Countries
; CURRENT APPLICATION NUMBER: US/10/219,051B
; CURRENT FILING DATE: 2003-05-09
```

```
RESULT 10
US-10-153-668-162
; Sequence 162, Application US/10153668
; GENERAL INFORMATION:
; APPLICANT: HONDA, Goichi
; APPLICANT: MATSUDA, Akio
; APPLICANT: MURAMATSU, Shuji
; APPLICANT: ISHIZAWA, Kenya
; TITLE OF INVENTION: STAT6 Activating Gene
; FILE REFERENCE: 1254-0207P
; CURRENT APPLICATION NUMBER: US/10/153,668
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 60/293,172
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 60/316,031
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/328,403
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: JP 2001-157043
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: JP 2001-260681
; PRIOR FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: JP 2001-313175
; PRIOR FILING DATE: 2001-10-10
; NUMBER OF SEQ ID NOS: 488
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 162
; LENGTH: 276
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-153-668-162

Alignment Scores:
Pred. No.:      2,52e-22      Length:      276
Score:          349.00        Matches:      76
Percent Similarity: 92.68%    Conservative: 0
Best Local Similarity: 92.68% Mismatches:      6
Query Match:    67.77%       Indels:       2
DB:             27           Gaps:         0

US-10-042-417A-29 (1-278) x US-10-153-668-162 (1-276)
QY      34 ATGAGCGCGTAGTGTTCGGCGAGTCCCG-GGNTCTCCGTAGACCGCGGAN-ACC 91
Db      1 MetGluProValGlyCysGlyGluCysArgGlySerValAspProArgSerThr 20
QY      92 TTCGTGTTAGTAACCTGCGGAGGTGGTGGAGCGTGTGCTACCTTCCTGCCGCCCAAG 151
Db      21 PheValLeuSerAsnLeuAlaGluValGluArgValLeuThrPheLeuProAlaLys 40
QY      152 GCGTTGCTCGGGTGGCTGCTGCTGCTATGGAGGGAGTGTGTCGCAGAGTATTG 211
Db      41 AlalaLeuArgValAlaCysValCysArgLeuTrpArgGluCysValArgValLeu 60
QY      212 CGGACCCATCGAGCGTAACCTGCATCTCCGAGCGCTGGCGAGCGCGCCACCTGGNG 271
Db      61 ArgThrHisArgSerValThrTrpIleSerAlaGlyLeuAlaGlyHisLeuGlu 80
QY      272 GGGCAT 277
Db      81 GlyHis 82

RESULT 11
US-10-219-051B-5836
; Sequence 5836, Application US/10219051B
; GENERAL INFORMATION:
; APPLICANT: The General Hospital Corporation doing business as Massachusetts General
; APPLICANT: Hospital / Bayer AG
; TITLE OF INVENTION: Nucleotide sequences involved in pain
; FILE REFERENCE: Lea 35693 Foreign Countries
; CURRENT APPLICATION NUMBER: US/10/219,051B
; CURRENT FILING DATE: 2003-05-09
```

```
; PRIOR APPLICATION NUMBER: US 60/312,147
; PRIOR FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: US 60/346,382
; PRIOR FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: US 60/333,347
; PRIOR FILING DATE: 2001-11-26
; NUMBER OF SEQ ID NOS: 14715
; SOFTWARE: Perl script
; SEQ ID NO 5836
; LENGTH: 277
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(277)
; OTHER INFORMATION: Xaa=Unknown amino acid
US-10-219-051B-5836

Alignment Scores:
Pred. No.: 2,52e-22 Length: 277
Score: 349.00 Matches: 76
Percent Similarity: 92.68% Conservativeness: 0
Best Local Similarity: 92.68% Mismatches: 6
Query Match: 67.77% Indels: 2
DB: 28 Gaps: 0

US-10-042-417A-29 (1-278) x US-10-219-051B-5836 (1-277)
QY 34 ATGAGCGCGTAGTGTCTCGCGGAGTCCCG-GGNTCTCTCGTAGACCGCGGAN-ACC 91
Db 1 MetGluProValGlyCysGlyGluCysArgGlySerValAspProArgSerThr 20
QY 92 TTCGTGTGTAGTAACCTGCGGAGGTGGTGGAGCGTGTGCTCACCTTCTCTGCCGCCAAG 151
Db 21 PheValLeuSerAsnLeuAlaGluValValGluArgValLeuThrPheLeuProAlaLys 40
QY 152 GCCTTGTCTGGGTGGCTCGCTCGCTTATGGAGGAGTGTGTGCGCAGAGATTG 211
Db 41 AlaLeuLeuArgValAlaLysValCysArgLeuTrpArgGlyCysValArgValLeu 60
QY 212 CGGACCATCGGAGCGTAACCTGGATCTCGCGAGGCTGGCGGAGCGCGCCACCTGGNG 271
Db 61 ArgThrHisArgSerValThrTrpIleSerAlaGlyLeuAlaGlyHisLeuGlu 80
QY 272 GGGCAT 277
Db 81 GlyHis 82

RESULT 12
US-10-219-051B-5839
; Sequence 5839, Application US/10219051B
; GENERAL INFORMATION:
; APPLICANT: The General Hospital Corporation doing business as Massachusetts General
; APPLICANT: Hospital / Bayer AG
; TITLE OF INVENTION: Nucleotide sequences involved in pain
; FILE REFERENCE: Lea 35693 Foreign Countries
; CURRENT APPLICATION NUMBER: US/10/219,051B
; CURRENT FILING DATE: 2003-05-09
; PRIOR APPLICATION NUMBER: US 60/312,147
; PRIOR FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: US 60/346,382
; PRIOR FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: US 60/333,347
; NUMBER OF SEQ ID NOS: 14715
; SOFTWARE: Perl script
; SEQ ID NO 5839
; LENGTH: 277
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(277)
; OTHER INFORMATION: Xaa=Unknown amino acid
US-10-219-051B-11414

Alignment Scores:
Pred. No.: 2,52e-22 Length: 277
Score: 349.00 Matches: 76
Percent Similarity: 92.68% Conservativeness: 0
Best Local Similarity: 92.68% Mismatches: 6
Query Match: 67.77% Indels: 2
DB: 28 Gaps: 0

US-10-042-417A-29 (1-278) x US-10-219-051B-11414 (1-277)
QY 34 ATGAGCGCGTAGTGTCTCGCGGAGTCCCG-GGNTCTCTCGTAGACCGCGGAN-ACC 91
Db 1 MetGluProValGlyCysGlyGluCysArgGlySerValAspProArgSerThr 20
QY 92 TTCGTGTGTAGTAACCTGCGGAGGTGGTGGAGCGTGTGCTCACCTTCTCTGCCGCCAAG 151
Db 21 PheValLeuSerAsnLeuAlaGluValValGluArgValLeuThrPheLeuProAlaLys 40
QY 152 GCCTTGTCTGGGTGGCTCGCTCGCTTATGGAGGAGTGTGTGCGCAGAGATTG 211
Db 41 AlaLeuLeuArgValAlaLysValCysArgLeuTrpArgGlyCysValArgValLeu 60
QY 212 CGGACCATCGGAGCGTAACCTGGATCTCGCGAGGCTGGCGGAGCGCGCCACCTGGNG 271
Db 61 ArgThrHisArgSerValThrTrpIleSerAlaGlyLeuAlaGlyHisLeuGlu 80
QY 272 GGGCAT 277
Db 81 GlyHis 82

RESULT 13
US-10-219-051B-11414
; Sequence 11414, Application US/10219051B
; GENERAL INFORMATION:
; APPLICANT: The General Hospital Corporation doing business as Massachusetts General
; APPLICANT: Hospital / Bayer AG
; TITLE OF INVENTION: Nucleotide sequences involved in pain
; FILE REFERENCE: Lea 35693 Foreign Countries
; CURRENT APPLICATION NUMBER: US/10/219,051B
; CURRENT FILING DATE: 2003-05-09
; PRIOR APPLICATION NUMBER: US 60/312,147
; PRIOR FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: US 60/346,382
; PRIOR FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: US 60/333,347
; NUMBER OF SEQ ID NOS: 14715
; SOFTWARE: Perl script
; SEQ ID NO 11414
; LENGTH: 277
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(277)
; OTHER INFORMATION: Xaa=Unknown amino acid
US-10-219-051B-11414

Alignment Scores:
Pred. No.: 2,52e-22 Length: 277
Score: 349.00 Matches: 76
Percent Similarity: 92.68% Conservativeness: 0
Best Local Similarity: 92.68% Mismatches: 6
Query Match: 67.77% Indels: 2
DB: 28 Gaps: 0

US-10-042-417A-29 (1-278) x US-10-219-051B-11414 (1-277)
QY 34 ATGAGCGCGTAGTGTCTCGCGGAGTCCCG-GGNTCTCTCGTAGACCGCGGAN-ACC 91
Db 1 MetGluProValGlyCysGlyGluCysArgGlySerValAspProArgSerThr 20
QY 92 TTCGTGTGTAGTAACCTGCGGAGGTGGTGGAGCGTGTGCTCACCTTCTCTGCCGCCAAG 151
Db 21 PheValLeuSerAsnLeuAlaGluValValGluArgValLeuThrPheLeuProAlaLys 40
QY 152 GCCTTGTCTGGGTGGCTCGCTCGCTTATGGAGGAGTGTGTGCGCAGAGATTG 211
Db 41 AlaLeuLeuArgValAlaLysValCysArgLeuTrpArgGlyCysValArgValLeu 60
QY 212 CGGACCATCGGAGCGTAACCTGGATCTCGCGAGGCTGGCGGAGCGCGCCACCTGGNG 271
Db 61 ArgThrHisArgSerValThrTrpIleSerAlaGlyLeuAlaGlyHisLeuGlu 80
QY 272 GGGCAT 277
Db 81 GlyHis 82
```

Db	41	AlaLeuLeuArgValAlaCysValCysArgLeuTrpArgGluCysValArgArgValLeu	60
Qy	212	CGAGCCCATCGGAGCGTAACCTGGATCTCCGAGGCGCTGGCGGAGCGCGCCACTCGNG	271
Db	61	ArgThrHisArgSerValThrTrpIleSerAlaGlyLeuAlaGluAlaGlyHisLeuGlu	80
Qy	272	GGGCAT	277
Db	81	GlyHis	32

APPLICANT: BAIRD, subject
APPLICANT: BAUGHN, Mariah R.
APPLICANT: BECHA, Shanya D.
APPLICANT: BOROWSKY, Mark L.
APPLICANT: BUFORD, Neil
APPLICANT: DING, bi
APPLICANT: ELLIOTT, Vicki S.
APPLICANT: EMERLING, Brooke M.
APPLICANT: GANDHI, Aameena R.
APPLICANT: GIETZEN, Kimberly J.
APPLICANT: GRIFFIN, Jennifer A.
APPLICANT: HAFALIA, April J.A.
APPLICANT: HONCHELL, Cynthia D.

; APPLICANT: EMERLING, Brooke M.
 ; APPLICANT: GANDHI, Ameena R.
 ; APPLICANT: GIETZEN, Kimberly J.
 ; APPLICANT: GRIFFIN, Jennifer A.
 ; APPLICANT: HAFALIA, April J.A.
 ; APPLICANT: HONCHELL, Cynthia D.
 ; APPLICANT: LAL, Preeti G.
 ; APPLICANT: LEE, Soo Yeun
 ; APPLICANT: LU, Dyung Aina M.
 ; APPLICANT: ARVIZU, Chandra S.
 ; APPLICANT: RAMKUNAR, Jayalaxmi
 ; APPLICANT: REDDY, Roopa
 ; APPLICANT: SANJANWALA, Madhu, M.

: APPLICANT:	LAL, Preeti G.
: APPLICANT:	LEE, Soo Yeun
: APPLICANT:	LU, Dyung Aina M.
: APPLICANT:	ARVIZU, Chandra S.
: APPLICANT:	RAMKUMAR, Jayalaxmi
: APPLICANT:	REDDY, Roopa
: APPLICANT:	SANJANWALA, Madhu, M.
: APPLICANT:	TANG, Y. Tom
: APPLICANT:	WALIA, Narinder K.
: APPLICANT:	WANG, Yu-mei, E.
: APPLICANT:	WARREN, Bridget A.
: APPLICANT:	XU, Yuning
: APPLICANT:	YANG, Cuming

: APPLICANT: TANG, Y. Tom
 : APPLICANT: WALIA, Narinder K.
 : APPLICANT: WANG, Yu-mei, E.
 : APPLICANT: WARREN, Bridget A.
 : APPLICANT: XU, Yuming
 : APPLICANT: YANG, Cunming
 : APPLICANT: YAO, Mcnique G.
 : APPLICANT: YUE, Henry
 : APPLICANT: ZEBARJADIAN, Yeganeh
 : TITLE OF INVENTION: PROTEINS ASSOCIATED WITH CELL GROWTH, DIFFERENTIATION, AND DEATH

: APPLICANT: YAO, Mcnique G.
 : APPLICANT: YUE, Henry
 : APPLICANT: ZEBARJADIAN, Yeganeh
 : TITLE OF INVENTION: PROTEINS ASSOCIATED WITH CELL GROWTH, DIFFERENTIATION, AND DEATH
 : FILE REFERENCE: PI-0417 PCT
 : CURRENT APPLICATION NUMBER: PCT/US02/11152
 : CURRENT FILING DATE: 2002-04-05
 : PRIOR APPLICATION NUMBER: 60/282,110; 60/283,294;
 : 60/291,562; 60/291,846; 60/293,727; 60/286,820; 60/287,228;
 : 60/291,562; 60/291,846; 60/293,727; 60/295,340;

```

, 60/295, 263; 60/349, 705;
,
, PRIOR FILING DATE: 2001-04-06; 2001-04-11; 2001-04-26; 2001-04-27;
,
, 2001-05-16; 2001-05-18; 2001-05-25; 2001-06-01;
,
, 2001-06-01; 2002-01-15
,
, NUMBER OF SEQ ID NOS: 42
,
, SOFTWARE: PERL Program
,
, SEQ ID NO 15
,
, LENGTH: 403
,
, TYPE: pep
,

```

```

1120: ENI
;
; ORGANISM: Homo sapiens
;
; FEATURE:
;
; NAME/KEY: misc_feature
;
; OTHER INFORMATION: Incyte ID No: 5565648CD1
PCT-US02-11152-15
;
; Alignment Scores:
Pred. No.: 2.49e-22 Length: 403

```

Score:	349.00	Matches:	76
Percent Similarity:	92.68%	Conservative:	0
Best Local Similarity:	92.68%	Mismatches:	6
Query Match:	67.77%	Indels:	2

```

DB:          1          Gaps:          0
US-10-042-417A-29 (1-278) x PCT-US02-11152-15 (1-403)
QY 34 ATGGAGCCGCTAGTGTCTTGGCGCGAGTCCCG-GGNTCCCTCCGTAGACCCCGCGAN-ACC 91
Db 1 MetGluProValGlyCysCysGlyGluCysArgGlySerSerValAspProArgSerThr 20
QY 92 TTCGTGTTGAGTAACCTGGCGGAGGTGGTGGAGCGTGTGCTACCTTCCTGCGCGCCAG 151
Db 21 PheValLeuSerAsnLeuAlaGluValValGluArgValLeuThrPheLeuProAlaLys 40
QY 152 GCGTGTGCTGGCGGTGGCTGCTGCCGCTTATGGAGGGAGTCTGTGCGCAGAGTATTG 211
Db 41 AlaLeuLeuArgValAlaCysValCysArgLeuTipArgGluCysValArgArgValLeu 60
QY 212 CGGACCCATCGGAGCGGTACCTGGATCTCGCGAGGCGCTGGCGAGGCGCGCCACCTGGNG 271
Db 61 ArgThrHisArgSerValThrTrpIleSerAlaGlyLeuAlaGluAlaGlyHisLeuGlu 80
QY 272 GGGCAT 277
Db 81 GlyHis 82

```

Search completed: August 28, 2003, 13:37:17
Job time : 237.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 28, 2003, 13:27:34 ; Search time 138.5 seconds
(without alignments)
549.103 Million cell updates/sec

Title: US-10-042-417A-29

Perfect score: 515

Sequence: 1 ccgtagtactgnttcggc.....cggccactgngggcatt 278

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 510680 seqs, 136781880 residues

Total number of hits satisfying chosen parameters: 1021360

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlh
-O=/cgn2_1/USPTO_spool/US10042417/runat_19082003_133542_9790/app_query.fasta_1.455
-DB=PublishedApplications_AA -QFWT=fastan -SUFFIX=rapb -MINMATCH=0.1
-LOOPEL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=BIOSUM62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10042417@cgn1.1.38 @runat_19082003_133542_9790
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PublishedApplications_AA:

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Score	Match Length	DB ID	Description

1	436.5	84.8	91	14	US-10-042-417-30	Sequence 30, Appl
2	349	67.8	276	15	US-10-153-668-162	Sequence 162, App
3	349	67.8	403	15	US-10-153-668-160	Sequence 160, App
4	229	44.5	44	14	US-10-042-417-63	Sequence 63, Appl
5	93.5	19.2	509	12	US-10-021-660-86	Sequence 86, Appl
6	91	18.6	680	15	US-10-177-293-59	Sequence 59, Appl
7	88.5	17.2	276	12	US-10-017-161-1576	Sequence 1576, A
8	88	18.0	558	15	US-10-156-761-10602	Sequence 10602, A
9	87.5	17.9	515	15	US-10-097-340-212	Sequence 212, App
10	87.5	17.9	515	15	US-10-171-311-156	Sequence 156, App
11	86	17.6	274	9	US-09-850-887-4	Sequence 4, Appli
12	85.5	17.5	193	9	US-09-864-761-37748	Sequence 37748, A
13	85	17.4	507	14	US-10-078-547-24	Sequence 24, Appl
14	84	17.2	1078	15	US-10-058-124-21	Sequence 21, Appl
15	84	17.2	1466	12	US-10-301-822-33	Sequence 33, Appl
16	84	17.2	1466	15	US-10-177-293-68	Sequence 68, Appl
17	83.5	17.1	984	12	US-10-017-161-1100	Sequence 1100, Ap
18	83	17.0	173	10	US-09-915-582-97	Sequence 97, Appl
19	83	17.0	609	15	US-10-156-761-7744	Sequence 7744, Ap
20	83	17.0	1199	15	US-10-128-714-8288	Sequence 8288, Ap
21	82.5	16.0	135	9	US-09-864-761-33360	Sequence 33360, A
22	82.5	16.0	1584	15	US-10-225-567A-342	Sequence 342, App
23	81.5	16.7	88	9	US-09-867-550-882	Sequence 882, App
24	81.5	15.8	189	12	US-10-199-672-534	Sequence 534, App
25	81.5	15.8	189	12	US-10-187-749-534	Sequence 534, App
26	81.5	15.8	189	12	US-10-194-457-534	Sequence 534, App
27	81.5	15.8	189	12	US-10-184-642-534	Sequence 534, App
28	81.5	15.8	189	14	US-10-052-586-534	Sequence 534, App
29	81.5	15.8	189	15	US-10-174-590-534	Sequence 534, App
30	81.5	15.8	189	15	US-10-176-758-534	Sequence 534, App
31	81.5	15.8	189	15	US-10-175-737-534	Sequence 534, App
32	81.5	15.8	189	15	US-10-173-706-534	Sequence 534, App
33	81.5	15.8	189	15	US-10-175-738-534	Sequence 534, App
34	81.5	15.8	189	15	US-10-175-752-534	Sequence 534, App
35	81.5	15.8	189	15	US-10-176-482-534	Sequence 534, App
36	81.5	15.8	189	15	US-10-176-757-534	Sequence 534, App
37	81.5	15.8	189	15	US-10-176-913-534	Sequence 534, App
38	81.5	15.8	189	15	US-10-180-552-534	Sequence 534, App
39	81.5	15.8	189	15	US-10-180-557-534	Sequence 534, App
40	81.5	15.8	189	15	US-10-173-700-534	Sequence 534, App
41	81.5	15.8	189	15	US-10-174-572-534	Sequence 534, App
42	81.5	15.8	189	15	US-10-174-579-534	Sequence 534, App
43	81.5	15.8	189	15	US-10-174-582-534	Sequence 534, App
44	81.5	15.8	189	15	US-10-174-588-534	Sequence 534, App
45	81.5	15.8	189	15	US-10-175-739-534	Sequence 534, App

ALIGNMENTS

RESULT 1

US-10-042-417-30
; Sequence 30, Application US/10042417
; Publication No. US2020123082A1
; GENERAL INFORMATION:
; APPLICANT: Pagano, M.
; TITLE OF INVENTION: METHODS TO IDENTIFY COMPOUNDS USEFUL FOR THE TREATMENT OF PROLIFERATIVE AND DIFFERENTIATIVE DISORDERS
; FILE REFERENCE: 5914-090-999
; CURRENT APPLICATION NUMBER: US/10/042,417
; CURRENT FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: 60/260,179
; PRIOR FILING DATE: 2001-01-05
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: all Xaa positions
; OTHER INFORMATION: Xaa-unknown amino acid residue
US-10-042-417-30

Alignment Scores:

Pred. No.: 1.56e-36 Length: 91
 Score: 436.50 Matches: 91
 Percent Similarity: 98.91% Conservative: 0
 Best Local Similarity: 98.91% Mismatches: 0
 Query Match: 84.76% Indels: 1
 DB: 14 Gaps: 1

US-10-042-417A-29 (1-278) x US-10-042-417-30 (1-91)

QY 2 CGTAGTACTGNTTCCGGCGGCTGTGAGAAATGGAGCCGGTAGNTGCTTGGCGCGAGT 61
 |||||
 Db 1 ArgSerThrGlyPheArgArgAlaGlyGluGluTrpSerArg-----LeuAlaAlaSer 19
 |||||
 QY 62 CCGGGNCTCCCTAGACCCCGGANACCTTCGTGTGAGTAACCTGGCGGAGGTGGTG 121
 |||||
 Db 20 ProGly**LeuArgArgProAla**ThrPheValLeuSerAsnLeuAlaGluValVal 39
 |||||
 QY 122 GAGCGTGTCTACCTTCCTCCCGCCCAAGCGCTTGTGCGGGTGGCGCTCGTGTGCCGC 181
 |||||
 Db 40 GluArgValLeuThrPheLeuProAlaLysAlaLeuLeuArgValAlaCysValCysArg 59
 |||||
 QY 182 TTATGAGGAGGTGTGCGCAGAGTATTCCGGACCCATCGGAGCGTAACCTGGATCTCC 241
 |||||
 Db 60 LeuTrpArgGluCysValArgArgValLeuArgThrHisArgSerValThrTrpIleSer 79
 |||||
 QY 242 GCAGCGCTGGCGGAGCGCCGCCACCTGGNGGGCAT 277
 |||||
 Db 80 AlaGlyLeuAlaGluAlaGlyHisLeu**GlyHis 91
 |||||

RESULT 2

US-10-153-668-162
 ; Sequence 162, Application US/10153668
 ; Publication No. US20030092616A1

GENERAL INFORMATION:

; APPLICANT: HONDA, Goichi
 ; APPLICANT: MATSUDA, Akio
 ; APPLICANT: MURAMATSU, Shuji
 ; APPLICANT: ISHIZAWA, Kenya
 ; TITLE OF INVENTION: STAT6 Activating Gene
 ; FILE REFERENCE: 1254-0207P
 ; CURRENT APPLICATION NUMBER: US/10/153,668
 ; CURRENT FILING DATE: 2002-05-24
 ; PRIOR FILING DATE: 2001-05-25
 ; PRIOR APPLICATION NUMBER: US 60/293,172
 ; PRIOR FILING DATE: 2001-08-30
 ; PRIOR APPLICATION NUMBER: US 60/316,031
 ; PRIOR FILING DATE: 2001-08-31
 ; PRIOR APPLICATION NUMBER: US 60/328,403
 ; PRIOR FILING DATE: 2001-10-12
 ; PRIOR APPLICATION NUMBER: JP 2001-157043
 ; PRIOR FILING DATE: 2001-05-25
 ; PRIOR APPLICATION NUMBER: JP 2001-260681
 ; PRIOR FILING DATE: 2001-08-30
 ; PRIOR APPLICATION NUMBER: JP 2001-313175
 ; PRIOR FILING DATE: 2001-10-10
 ; NUMBER OF SEQ ID NOS: 488
 ; SOFTWARE: PatentIn ver. 2.0
 ; SEQ ID NO 162
 ; LENGTH: 276
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-153-668-162

Alignment Scores:

Pred. No.: 1.62e-27 Length: 276
 Score: 349.00 Matches: 76
 Percent Similarity: 92.68% Conservative: 0
 Best Local Similarity: 92.68% Mismatches: 6
 Query Match: 84.77% Indels: 2
 DB: 15 Gaps: 0

US-10-042-417A-29 (1-278) x US-10-153-668-162 (1-276)

QY 34 ATGAGCGCGGTAGNTGCTTGGCGGAGTCCCG-GGNTCCTCCGTAGACCCCGCGGAN-ACC 91
 |||||
 Db 1 MetGluProValGlyCysGlyCysArgGlySerSerValAspProArgSerThr 20
 |||||
 QY 92 TTCGTGTGAGTAACCTGGCGGAGGTGGTGAGCGGTGTGCTACCTTCTCCCGCCCAAG 151
 |||||
 Db 21 PheValLeuSerAsnLeuAlaGluValValArgValLeuThrPheLeuProAlaLys 40
 |||||
 QY 152 GCCTTGTGCGGGTGGCTCGCTGTCGGCTTATGGAGGAGTGTGTGCGCAGAGTATTG 211
 |||||
 Db 41 AlaLeuLeuArgValAlaCysValCysArgLeuTrpArgGluCysValArgValLeu 60
 |||||
 QY 212 CGAGCCCATCGGACCGTAACTCGCATCTCGCGAGCGCTGGCGGAGCGCCGCCACCTCGNG 271
 |||||
 Db 61 ArgThrHisArgSerValThrTrpIleSerAlaGlyLeuAlaGluAlaGlyHisLeuGlu 80
 |||||
 QY 272 GGGCAT 277
 |||||
 Db 81 GlyHis 82
 |||||

RESULT 3

US-10-153-668-160
 ; Sequence 160, Application US/10153668
 ; Publication No. US20030092616A1

GENERAL INFORMATION:

; APPLICANT: HONDA, Goichi
 ; APPLICANT: MATSUDA, Akio
 ; APPLICANT: MURAMATSU, Shuji
 ; APPLICANT: ISHIZAWA, Kenya
 ; TITLE OF INVENTION: STAT6 Activating Gene
 ; FILE REFERENCE: 1254-0207P
 ; CURRENT APPLICATION NUMBER: US/10/153,668
 ; CURRENT FILING DATE: 2002-05-24
 ; PRIOR FILING DATE: 2001-05-25
 ; PRIOR APPLICATION NUMBER: US 60/293,172
 ; PRIOR FILING DATE: 2001-08-30
 ; PRIOR APPLICATION NUMBER: US 60/316,031
 ; PRIOR FILING DATE: 2001-08-31
 ; PRIOR APPLICATION NUMBER: US 60/328,403
 ; PRIOR FILING DATE: 2001-10-12
 ; PRIOR APPLICATION NUMBER: JP 2001-157043
 ; PRIOR FILING DATE: 2001-05-25
 ; PRIOR APPLICATION NUMBER: JP 2001-260681
 ; PRIOR FILING DATE: 2001-08-30
 ; PRIOR APPLICATION NUMBER: JP 2001-313175
 ; PRIOR FILING DATE: 2001-10-10
 ; NUMBER OF SEQ ID NOS: 488
 ; SOFTWARE: PatentIn ver. 2.0
 ; SEQ ID NO 160
 ; LENGTH: 403
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-153-668-160

Alignment Scores:

Pred. No.: 1.68e-27 Length: 403
 Score: 349.00 Matches: 76
 Percent Similarity: 92.68% Conservative: 0
 Best Local Similarity: 92.68% Mismatches: 6
 Query Match: 67.77% Indels: 2
 DB: 15 Gaps: 0

US-10-042-417A-29 (1-278) x US-10-153-668-160 (1-403)

QY 34 ATGAGCGCGGTAGNTGCTTGGCGGAGTCCCG-GGNTCCTCCGTAGACCCCGCGGAN-ACC 91
 |||||
 Db 1 MetGluProValGlyCysGlyCysArgGlySerSerValAspProArgSerThr 20
 |||||
 QY 92 TTCGTGTGAGTAACCTGGCGGAGGTGGTGAGCGGTGTGCTACCTTCTCCCGCCCAAG 151
 |||||
 Db 21 PheValLeuSerAsnLeuAlaGluValValArgValLeuThrPheLeuProAlaLys 40
 |||||
 QY 152 GCCTTGTGCGGGTGGCTCGCTGTCGGCTTATGGAGGAGTGTGTGCGCAGAGTATTG 211
 |||||

```

Db 41 AlaLeuArgValAlaCysValCysArgLeuTriPArgGluCysValArgArgValLeu 60
QY 212 CGGACCATCGGAGCGTAACCTGGATCTCGCAGGCTGCGGAGCGCGCCACCTGGNG 271
Db 61 ArgThrHisArgSerValThrTrpIleSerAlaGlyLeuAlaGluAlaGlyHisLeuGlu 80
QY 272 GGCAT 277
Db 81 GlyHis 82

RESULT 4
US-10-042-417-63
; Sequence 63, Application US/10042417
; Publication No. US20020123082A1
; GENERAL INFORMATION:
; APPLICANT: Pagano, M.
; TITLE OF INVENTION: METHODS TO IDENTIFY COMPOUNDS USEFUL FOR THE TREATMENT OF
; TITLE OF INVENTION: PROLIFERATIVE AND DIFFERENTIATIVE DISORDERS
; FILE REFERENCE: 5914-090-999
; CURRENT APPLICATION NUMBER: US/10/042,417
; CURRENT FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: 60/260,179
; PRIOR FILING DATE: 2001-01-5
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 63
; LENGTH: 44
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-042-417-63

Alignment Scores:
Pred. No.: 2,7e-15 Length: 44
Score: 229.00 Matches: 44
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Watch: 44.47% Indels: 0
DB: 14 Gaps: 0

US-10-042-417a-29 (1-278) x US-10-042-417-63 (1-44)

QY 107 CTGCGGAGGTGTGGAGCTGTGCTCCTCTGCGCCAGCGTGTGCTGCGGGT 166
Db 1 LeuAlaGluValGluArgValLeuThrPheLeuProAlaLysAlaLeuArgVal 20
QY 167 GCTGCGGTGCGCCCTATGAGGAGGTGTGCGCAGAGTATGCGGACCCATCGGAGC 226
Db 21 AlaCysValCysArgLeuTriPArgGluCysValArgArgValLeuArgThrHisArgSer 40
QY 227 GTAACCTGGATC 238
Db 41 ValThrTrpIle 44

RESULT 5
US-10-021-660-86
; Sequence 86, Application US/10021660
; Publication No. US20030152926A1
; GENERAL INFORMATION:
; APPLICANT: Murray, Richard
; APPLICANT: Glynn, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: EOS Biotechnology, Inc.
; TITLE OF INVENTION: No. US20030152926A1 Methods of Diagnosis of Angiogenesis,
; TITLE OF INVENTION: Compositions and Methods of Screening for Angiogenesis
; TITLE OF INVENTION: Modulators
; FILE REFERENCE: 018501-0007100S
; CURRENT APPLICATION NUMBER: US/10/021,660
; CURRENT FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: US/09/784,356
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/637,977
; PRIOR FILING DATE: 2000-08-11

```

```

; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 86
; LENGTH: 509
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-021-660-86

Alignment Scores:
Pred. No.: 0.265 Length: 509
Score: 93.50 Matches: 47
Percent Similarity: 41.27% Conservative: 5
Best Local Similarity: 37.30% Mismatches: 31
Query Watch: 19.16% Indels: 43
DB: 12 Gaps: 9

US-10-042-417a-29 (1-278) x US-10-021-660-86 (1-509)

QY 272 CCNOCAGTGGCGGCTCGCCAGGCTGCGGAGATCCAGGTACGTCGATGGGTCC 213
Db 151 ProProAlaProProProProProProProProProProProProProProPro 168
QY 212 GCAATACTTGGCGCA---CACACT----- 192
Db 169 AlaLeuProCysAlaGlyHisThrArgArgArgArgArgArgArgArgArgSer 188
QY 191 CCTCCATAGC-----GGCACACGACGCGCCACCC--- 162
Db 189 ProProLeuSerGlyProProGlyArgAlaSerProArgGlyAlaArgProProLeu 208
QY 161 -----GCAGCACGCTTGG---CGGSCAGGAAGGTGAGCACACGCT----- 123
Db 209 LeuArgAlaAlaProThrProSerProArgAlaLeuAlaProAlaAlaSerProPro 228
QY 122 ---CCACCCTCGCCAGGTACTTACTCAACAGCAA-----GGTNT 87
Db 229 ProProProProProProProProProProProProProProProProProGlyse 248
QY 86 CCGCGGCTTACGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 30
Db 248 rSerGlySerThrGlnThrSerGlyAla-AlaAlaAlaValAlaAlaLeuGlySerS 268
QY 29 CACGAGCCGCGCG 16
Db 268 erProGlyArgArg 272

RESULT 6
US-10-177-293-59
; Sequence 59, Application US/10177293
; Publication No. US20030124128A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Zhao, Xumei
; APPLICANT: Gannavapu, Manjula
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Mertens, Maureen
; APPLICANT: Myer, Vic
; APPLICANT: Wang, Youzhen
; APPLICANT: Xu, Yongyao
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Monahan, John
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Bast Jr., Robert C.
; APPLICANT: Hortobagyi, Gabriel N.
; APPLICANT: Pusztai, Lajos
; APPLICANT: Meric, Funda
; APPLICANT: Sahin, Aysegul
; APPLICANT: Mills, Gordon B.
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT
; TITLE OF INVENTION: PREVENTION, AND THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-038
; CURRENT APPLICATION NUMBER: US/10/177,293

```

Alignment Scores:				
Pred. No.:	0.813	Length:	276	
Score:	88.50	Matches:	29	
Percent Similarity:	41.46%	Conservative:	5	
Best Local Similarity:	35.37%	Mismatches:	28	
Query Match:	17.18%	Indels:	20	
DB:	12	Gaps:	3	
US-10-042-417A-29 (1-278) x US-10-017-161-1576 (1-276)				
Qy	60	GTCCCGGGNCTCCGTAGACCCGCGANACCTTCGTCTGAGTAACCTGGCGAGGTGG	119	
Db	117	ValProSerThrCysLeuSerAlaAlaProAlaCys-----	129	
Qy	120	TGGAGCGTGTCTACCTTCCTGCGCCGCGACAGGCTTCTCGGGTGGCCCTGCG-	173	
Db	130	-----CysCysThrSerSerAlaThrSerArgArgCysCysThrTrpProAlaAlaArgLy	147	
Qy	174	-----TGTCGCCCTTATCGAGGGAGTGTGTCGCAGAGTATTGCGGACCCATCGGA	224	
Db	148	LeuArgArgArgAlaAlaSerAlaGlyArgLeuProGlyAlaAlaAlaAlaLeuGly	167	
Qy	225	GCGTAACCTGGATCTC-----CGCAGGCTGCGGAGCGCGCCACCTGG	269	
Db	168	AlaHisProGly-LeuLeuTrpArgHisArgArgProGlyArgLeuArgCysProLy	187	
Qy	270	NGGG	273	
Db	187	sGly	188	
RESULT 8				
US-10-156-761-10602				
; Sequence 10602, Application US/10156761				
; Publication No. US20030119018A1				
; GENERAL INFORMATION:				
; APPLICANT: OMURA, SATOSHI				
; APPLICANT: IKEDA, HARUO				
; APPLICANT: ISHIKAWA, JUN				
; APPLICANT: HORIKAWA, HIROSHI				
; APPLICANT: SHIBA, TADAYOSHI				
; APPLICANT: SAKAKI, YOSHIYUKI				
; APPLICANT: HATTORI, MASAHIRA				
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES				
; FILE REFERENCE: 249-262				
; CURRENT APPLICATION NUMBER: US/10/156,761				
; CURRENT FILING DATE: 2002-05-29				
; PRIOR APPLICATION NUMBER: JP 2001-204089				
; PRIOR FILING DATE: 2001-05-30				
; PRIOR APPLICATION NUMBER: JP 2001-272697				
; PRIOR FILING DATE: 2001-08-02				
; NUMBER OF SEQ ID NOS: 15109				
; SEQ ID NO 10602				
; LENGTH: 558				
; TYPE: PRT				
; ORGANISM: Streptomyces avermitilis				
US-10-156-761-10602				
Alignment Scores:				
Pred. No.:	0.98	Length:	558	
Score:	88.00	Matches:	32	
Percent Similarity:	37.50%	Conservative:	7	
Best Local Similarity:	30.77%	Mismatches:	45	
Query Match:	18.03%	Indels:	20	
DB:	15	Gaps:	6	
US-10-042-417A-29 (1-278) x US-10-156-761-10602 (1-558)				
Qy	272	CNCCAGGTGGCGGCTCCGCCAGGC---	CTGCGGAGATCCAGGTACCTCGGATGG	216
Db	221	ProProGlyProThrProProProAlaProAlaProAlaProAlaProAlaPro	240	
Qy	215	TCCGCAATACTTCGGCCACACACTCCCTCCCATAGCGGCACACGACGCGCCACGCCGACGA	156	

DB: 15 Gaps: 5

US-10-042-417A-29 (1-278) x US-10-097-340-212 (1-515)

269 CCAGGTGGCGGCTCCGCGAGGCTCCGGAGATCCAGGTTACGCTCCGATGGTCCGCA 210
 ||| :||| ||| |||||
 122 ProAspAsnIysProAlaPro-----GlySerThr 131
 209 ATACTCTCGCACAC---ACTCCCTCATAGCGGCACACGACGAGGCCACCCGACGCAACG 153
 ||| ||| ||| ||||| ||||| :|||
 132 AlaProProAlaHisGlyValThrSerAlaProAspThrArgProProGlySerThr 151
 152 CTTGGCGCGCAGGAAGGTGAGCAGCGCTCCACCA----- 117
 |||||
 152 -----AlaProProAlaHisGlyValThrSerAlaPro 162
 116 -----CCTCGCGCAGGTACTCAACACGAGGTTTCGGGGGTCTACGAGGAGNC 66'
 ||||| ||||| :||| ||| |||
 163 AspThrArgProProGlySerThrAla-ProAlaAlaHisGlyValThrSerAlaPro 182
 ||||| ||||| ||||| ||||| :|||
 65 CGGACTCTCGCGCAGCANCTACCGGTCTCCATTCTCACCAGCGCCGCGGAANCCAGTAC 6
 ||||| ||||| ||||| ||||| :|||
 182 oAspThrArg-----ProAlaProGlySerThrAlaProProAlaHisGlyVal 198
 5 TACG 2
 |||
 198 lThr 199

RESULT 10

US-10-171-311-156

; Sequence 156, Application US/10171311

; Publication No. US20030087270A1

; GENERAL INFORMATION:

; APPLICANT: Schlegel, Robert

; APPLICANT: Chen, Yan

; APPLICANT: Zhao, Ximei

; APPLICANT: Monahan, John

; APPLICANT: Kamatkar, Shubhangi

; APPLICANT: Glatt, Karen

; APPLICANT: Gannavakapu, Manjula

; APPLICANT: Hoerish, Sebastian

; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR

; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY

; TITLE OF INVENTION: OF CERVICAL CANCER

; FILE REFERENCE: MRI-035

; CURRENT APPLICATION NUMBER: US/10/171,311

; CURRENT FILING DATE: 2002-06-12

; PRIOR APPLICATION NUMBER: US 60/298,159

; PRIOR FILING DATE: 2001-06-13

; PRIOR APPLICATION NUMBER: US 60/298,155

; PRIOR FILING DATE: 2001-06-13

; PRIOR APPLICATION NUMBER: US 60/335,936

; PRIOR FILING DATE: 2001-11-14

; NUMBER OF SEQ ID NOS: 238

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 156

; LENGTH: 515

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-171-311-156

Alignment Scores:

Pred. No.:	1.09	Length:	515
Score:	87.50	Matches:	32
Percent Similarity:	35.29%	Conservative:	4
Best Local Similarity:	31.37%	Mismatches:	30
Query Match:	17.93%	Indels:	36
DB:	15	Gaps:	5

US-10-042-417A-29 (1-278) x US-10-171-311-156 (1-515)

269 CCAGGTGGCGGCTCCGCGAGGCTCCGGAGATCCAGGTTACGCTCCGATGGTCCGCA 210
 ||| :||| ||| |||||

Db 122 ProAspAsnLysProAlaPro-----GlySerThr 131
 Qy 209 ATACTCTGGGCACAC-----ACTCCCTCCATAGCGGCACAGCGAGCCACCGCAGCAACG 153
 Db 132 AlaProProAlaHisGlyValThrSerAlaProAspThrArgProProGlySerThr 151
 Qy 152 CCTTGGCGGCGAGGAGGTGAGCAGACGCTCCACCA----- 117
 Db 152 -----AlaProProAlaHisGlyValThrSerAlaPro 162
 Qy 116 -----CCTCGCCAGGTACTCAACACGAGGTTNTCCGCGGGTCTACGGAGGANC 66
 Db 163 AspThrArgProProGlySerThrAla-ProAlaAlaHisGlyValThrSerAlaPr 182
 Qy 65 CGGACTCGCGCAGCAGCTACCGGTCCATTCCTCACCAGCCCGCCGGAANCAAGTAC 6
 Db 182 OAspThrArg-----ProAlaProGlySerThrAlaProProAlaHisGlyVa 198
 Qy 5 TACG 2
 Db 198 lThr 199
 RESULT 11
 ; Sequence 4, Application US/09850887
 ; Patent No. US20020009778A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lal, Preeti
 ; Corley, Neil C.
 ; Gorgone, Gina
 ; TITLE OF INVENTION: THYROID AND PITUITARY MEMBRANE PROTEIN
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
 ; STREET: 3174 PORTER DRIVE
 ; CITY: PALO ALTO
 ; STATE: CALIFORNIA
 ; COUNTRY: USA
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/850,887
 ; FILING DATE: 07-May-2001
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 09/087,678
 ; FILING DATE: <Unknown>
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: CERRONE, MICHAEL C.
 ; REGISTRATION NUMBER: 39,132
 ; REFERENCE/DOCKET NUMBER: PF-0535 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (650) 855-0555
 ; TELEFAX: (650) 845-4166
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 274 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; LIBRARY: GenBank
 ; CLONE: g206712
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
 US-09-850-887-4
 Alignment Scores: 1.47 Length: 274
 Pred. No.: 86.00 Matches: 35
 Score:

Percent Similarity: 38.78% Conservative: 3
 Best Local Similarity: 35.71% Mismatches: 31
 Query Match: 17.62% Indels: 29
 DB: 9 Gaps: 5
 US-10-042-417A-29 (1-278) x US-09-850-887-4 (1-274)
 Qy 272 CCNCCAGTGGCGGGCTCCGCCAGGCTCCGGAGATCCAGTTACGCTCCGATGGGTCC 213
 Db 60 ProProGlnGlyProProProProGly-----Gly-Pr 70
 Qy 212 GCAATACTCTGCGCACACACTCCTCCATTAAGCGGCACACGACGAGCCACCCGCGCAACG 153
 Db 70 oGlnGlnLysProGlnGlyProProPro-----ProGlyGlyProGlnGlnAr 86
 Qy 152 CATT-----GCGCGGCAGGAGGTGACACACACGCTCCACCACT 114
 Db 86 gProGlnGlyProProProProGlyProGlnGlnGly--ProGlnGlyProProPro 105
 Qy 113 CGCCAGGTACTCAACACGA---AGGTNTCCGCGGTCTACGGAGGAGNCCCGGGACTCG 57
 Db 106 ProGlyGlyProGlnGlnArgProGlnGlyProProProProGlyGlyProGlnGlnGly 125
 Qy 56 CGCAGCAGCTACCGGCTCCATTCCTCACCAGCCCGCGGAGNCCAG 9
 Db 126 ProGlnGly-----ProProProProGlyGlyProGln 136
 RESULT 12
 US-09-864-761-37748
 ; Sequence 37748, Application US/09864761
 ; Patent No. US20020048763A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Penn, Sharron G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.
 ; APPLICANT: Chen, Wensheng
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
 ; FILE REFERENCE: Aemica-X-1
 ; CURRENT APPLICATION NUMBER: US/09/864,761
 ; CURRENT FILING DATE: 2001-05-23
 ; PRIOR APPLICATION NUMBER: US 60/180,312
 ; PRIOR FILING DATE: 2000-02-04
 ; PRIOR APPLICATION NUMBER: US 60/207,456
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: US 09/632,366
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR APPLICATION NUMBER: GB 24263.6
 ; PRIOR FILING DATE: 2000-10-04
 ; PRIOR APPLICATION NUMBER: US 60/236,359
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: PCT/US01/00666
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00667
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00664
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00669
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00665
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00668
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00663
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00662
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00661
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00670
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: US 60/234,687
 ; PRIOR FILING DATE: 2000-09-21

; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 37748
; LENGTH: 193
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC002467.1
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.1
; OTHER INFORMATION: EST_HUMAN HIT: AA831893.1, EVALUATE 7.00e-07
US-09-864-761-37748

Alignment Scores:
Pred. No.: 1.59 Length: 193
Score: 85.50 Matches: 33
Percent Similarity: 39.00% Conservative: 6
Best Local Similarity: 33.00% Mismatches: 35
Query Match: 17.52% Indels: 27
DB: 9 Gaps: 5

US-10-042-417A-29 (1-278) x US-09-864-761-37748 (1-193)

QY 272 CCNCCAGTGGCCGGCTCCGCCAGGCTCGGAGATCCAGGTACGTCGGATGGTCC 213
Db 105 ProProProProProProProProProProGlnHisGlyGlyProValThrAlaProProProH 124
QY 212 GCAATACTCTCGGCACACACTCCCTCCGATA-----AGCGGCACACGAGCCACCC 163
Db 124 SHisTyAsnProAsnSerLeuProGlnPheThrGluAspGlnGlyThrLeuSerProPr 144
QY 162 CGCAGCACGGCT-----TGGCGGCACGAGAGGTGAGCACACGC 124
Db 144 oPheThrGlnProGlyGlyMetSerProGlyIleTrpProAlaProArg-----G 161
QY 123 TCCACCACTCCGCCAGGTACTCAACACGAGGTNTCCGGCGTCTACGAGGAGCCG 64
Db 161 yProProPro-----ProArgLeuGlnGlyProProSe 173
QY 63 GGACTCGCGCAGCAGCAGTACCGGCTCCATTCCTCCACGAGCCCGCGGAANCCAGTAC 6
Db 173 rGlnThrProLeuProGly-----ProHisProAspGlnThrArgTyr 188

RESULT 13
US-10-078-547-24
; Sequence 24, Application US/10078547
; Publication No. US20020199211A1
; GENERAL INFORMATION:
; APPLICANT: Narayanasamy Ramesh
; APPLICANT: Miguel A. de la Fuente
; APPLICANT: Ines M. Anton
; APPLICANT: Ralf S. Geha
; TITLE OF INVENTION: WIP, A WASP-Associated Protein
; FILE REFERENCE: 1242.1022-005
; CURRENT APPLICATION NUMBER: US/10/078,547
; CURRENT FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: 09/599,287
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: PCT/US98/27501
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: 60/101,457
; PRIOR FILING DATE: 1998-09-23

; PRIOR APPLICATION NUMBER: 60/068,533
; PRIOR FILING DATE: 1997-12-23
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 507
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Translated WIP ORF No. US20020199211A1 3
US-10-078-547-24

Alignment Scores:
Pred. No.: 1.97 Length: 507
Score: 85.00 Matches: 31
Percent Similarity: 45.88% Conservative: 8
Best Local Similarity: 36.47% Mismatches: 34
Query Match: 17.42% Indels: 12
DB: 14 Gaps: 5

US-10-042-417A-29 (1-378) x US-10-078-547-24 (1-507)

QY 272 CCNCCAGTGGCCGGCT---CCGCCAGGCTCGGAGATCCAGGTACGTCGGATGGG 216
Db 314 ProProProSerArgProGlyProProProProProProProSerSerSerGlyAsnAspGlu 333
QY 215 TCCGCAAACTCTCGGCACACACTCCCTCCATAAGCGGCACACGAGCCACCCGACCA 156
Db 334 ThrProArgLeuProGlnArgAsnLeuSerLeuSerSerSer---ThrProProLeuPro 352
QY 155 ACGCCTTTCGGCGGCAGGAAGGTGAGCACACGCTCCACCCACCTCCGCCAGGTACTCAACA 96
Db 353 SerProGlyArgSerGly-----ProLeuProProProValProSerGlu 367
QY 95 CGAAGGTHTCGGCGGTCTACGAGGAGNCCCGGACTCGCGCAGCAGCAGTACCGGCTCC 36
Db 368 ArgProProProProVal-----ArgAsp---ProProGlyArgSerGlyPro 382
QY 35 ATTCTCTCACACGCC 21
Db 383 LeuProProProPro 387

RESULT 14
US-10-058-124-21
; Sequence 21, Application US/10058124
; Publication No. US20030119058A1
; GENERAL INFORMATION:
; APPLICANT: Qvist, Per
; Bcnde, Martin

TITLE OF INVENTION: A Method for Assaying Collagen Fragments
in Body Fluids, A Test Kit and Means for Carrying Out th
Method and Use of the Method to Diagnose the Presence of
Disorders Associated with the Metabolism of

NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Darby & Darby PC
STREET: 805 Third Avenue
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/058,124
FILING DATE: 29-Jan-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/570,573
FILING DATE: 2002-MAY-12


```
; APPLICATION NUMBER: 08/187,319
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Gogoris, Adda C.
; REGISTRATION NUMBER: 29,714
; REFERENCE/DOCKET NUMBER: 4305/08701
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-527-7700
; TELEFAX: 212-753-6237
; TELEX: 236687
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1078 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; IMMEDIATE SOURCE:
; CLONE: COLLAGEN ALPHA 1 (III)
; SEQUENCE DESCRIPTION: SEQ ID NO: 21:
US-10-058-124-21

Alignment Scores:
Pred. No.: 2.69 Length: 1078
Score: 84.00 Matches: 31
Percent Similarity: 33.04% Conservative: 6
Best Local Similarity: 27.68% Mismatches: 31
Query Match: 17.21% Indels: 44
DB: 15 Gaps: 5

US-10-042-417A-29 (1-278) x US-10-058-124-21 (1-1078)
QY 269 CCAGTGGCGGCTCCGCCAGGCTCGGAGATCCAGGTACCTCCGATGGGTCCGCA 210
Db 533 ProGlyGluArgGlyProGlyLeu-
QY 209 ATACTCTGCGCACACACTCCCTCCATAAGCGGCACACGAGCGCAGGCGCAGCAACGCCT 150
Db 542 -----AlaGlyAlaProGlyLeuArgGlyGlyAlaGlyProGlyPro----- 556
QY 149 TGGCGGCGCAGGAGGTGAGCAGCAGCTCCACACCTCCGCCAGGTACTCA----- 99
Db 557 ---GluGlyGlyGlyAlaAlaGlyProGlyProGlyAlaAlaGlyThrPro 575
QY 98 ---ACAGCAAGGTTCCGGGCTACGAGGA----- 69
Db 576 GlyLeuGlnGlyMetProGlyGluArgGlyGlyLeuGlySerProGlyProLysGlyAsp 595
QY 68 -----NCCGGGACTCGCGCAGCA 48
Db 596 LysGlyGluProGlyGlyProGlyAlaAspGlyValProGlyLysAspGlyProArgGly 615
QY 47 NCTACCGGTCCATT---CCTACACGCGCGCGGA 15
Db 616 ProThrGlyProLysGlyProGlyProGlyProAlaGly 627

RESULT 15
US-10-301-822-33
; Sequence 33, Application US/10301822
; Publication No. US20030148410A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: Burgart, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF COLON CANCER
```

```
; FILE REFERENCE: MPM01-029p2RNM
; CURRENT APPLICATION NUMBER: US/10/301,822
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 33
; LENGTH: 1466
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-301-822-33

Alignment Scores:
Pred. No.: 2.77 Length: 1466
Score: 84.00 Matches: 31
Percent Similarity: 33.04% Conservative: 6
Best Local Similarity: 27.68% Mismatches: 31
Query Match: 17.21% Indels: 44
DB: 12 Gaps: 5

US-10-042-417A-29 (1-278) x US-10-301-822-33 (1-1466)
QY 269 CCAGTGGCGGCTCCGCCAGGCTCGGAGATCCAGGTACCTCCGATGGGTCCGCA 210
Db 680 ProGlyGluArgGlyProGlyLeu-
QY 209 ATACTCTGCGCACACACTCCCTCCATAAGCGGCACACGAGCGCAGCAGCAACGCCT 150
Db 689 -----AlaGlyAlaProGlyLeuArgGlyGlyAlaGlyProGlyPro----- 703
QY 149 TGGCGGCGCAGGAGGTGAGCAGCAGCTCCACACCTCCGCCAGGTACTCA----- 99
Db 704 ---GluGlyGlyGlyAlaAlaGlyProGlyProGlyProGlyAlaAlaGlyThrPro 722
QY 98 ---ACAGCAAGGTTCCGGGCTACGAGGA----- 69
Db 723 GlyLeuGlnGlyMetProGlyGluArgGlyGlyLeuGlySerProGlyProLysGlyAsp 742
QY 68 -----NCCGGGACTCGCGCAGCA 48
Db 743 LysGlyGluProGlyGlyProGlyAlaAspGlyValProGlyLysAspGlyProArgGly 762
QY 47 NCTACCGGTCCATT---CCTACACGCGCGCGGA 15
Db 763 ProThrGlyProLysGlyProGlyProGlyProAlaGly 774

Search completed: August 28, 2003, 13:42:48
Job time : 143.5 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 28, 2003, 13:24:34 ; Search time 19.5 Seconds

(without alignments)
1206.402 Million cell updates/sec

Title: US-10-042-417A-29

Perfect score: 515

Sequence: 1 cctagtactggttcggc.....cgsgccacctgnggggcatc 278

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Deiop 6.0 , Delext 7.0

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 657434

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10042417/runat_19082003_133539_9668/app_query.fasta_1.455
-DB=Issued Patents AA -QFMM=fastan -SUFFIX=rai -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10042417 -CGN_1_1_18_erunat_19082003_133539_9668 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: /cgn2_6/ptodata/2/iaa/5A.COMB.pep.*
- 2: /cgn2_6/ptodata/2/iaa/5B.COMB.pep.*
- 3: /cgn2_6/ptodata/2/iaa/6A.COMB.pep.*
- 4: /cgn2_6/ptodata/2/iaa/6B.COMB.pep.*
- 5: /cgn2_6/ptodata/2/iaa/PCTUS.COMB.pep.*
- 6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	101	20.7	162	4	US-09-252-991A-32101 Sequence 32101, A
C 2	100.5	19.5	328	4	US-09-252-991A-26586 Sequence 26586, A
C 3	97.5	20.0	258	4	US-09-252-991A-16884 Sequence 16884, A
C 4	96.5	19.8	161	4	US-09-252-991A-32112 Sequence 32112, A
C 5	95.5	19.6	142	4	US-09-252-991A-31831 Sequence 31831, A
C 6	94	19.3	144	4	US-09-252-991A-26492 Sequence 26492, A
C 7	94	19.3	478	4	US-09-252-991A-20429 Sequence 20429, A
C 8	93.5	18.2	554	4	US-09-252-991A-28232 Sequence 28232, A
C 9	93.5	19.2	618	4	US-09-252-991A-28358 Sequence 28358, A
C 10	92.5	19.0	165	4	US-09-252-991A-19409 Sequence 19409, A
C 11	92.5	19.0	457	4	US-09-252-991A-22682 Sequence 22682, A
C 12	92.5	19.0	574	4	US-09-252-991A-30868 Sequence 30868, A

c 13	92	18.9	256	4	US-09-252-991A-30033	Sequence 30033, A
c 14	92	18.9	292	4	US-09-252-991A-20173	Sequence 20173, A
c 15	91	18.6	145	4	US-09-252-991A-21532	Sequence 21532, A
c 16	91	18.6	299	4	US-09-252-991A-21789	Sequence 21789, A
c 17	90.5	18.5	221	4	US-09-252-991A-19403	Sequence 19403, A
c 18	90.5	18.5	262	4	US-08-311-731A-333	Sequence 333, App
c 19	90.5	18.5	347	4	US-09-252-991A-32763	Sequence 32763, A
c 20	90.5	18.5	370	4	US-09-252-991A-18438	Sequence 18438, A
c 21	90.5	18.5	663	4	US-09-252-991A-30843	Sequence 30843, A
c 22	90.5	18.5	977	4	US-09-252-991A-16655	Sequence 16655, A
c 23	90	18.4	343	4	US-09-252-991A-24630	Sequence 24630, A
c 24	90	18.4	632	4	US-09-252-991A-25344	Sequence 25344, A
c 25	89.5	18.3	289	4	US-09-252-991A-25496	Sequence 25496, A
c 26	89.5	18.3	322	4	US-09-252-991A-24076	Sequence 24076, A
c 27	89	17.3	147	4	US-09-252-991A-26082	Sequence 26082, A
c 28	89	18.2	248	4	US-09-252-991A-30679	Sequence 30679, A
c 29	89	18.2	491	4	US-09-252-991A-19396	Sequence 19396, A
c 30	88.5	17.2	146	4	US-09-252-991A-32900	Sequence 32900, A
c 31	88.5	17.2	286	4	US-09-252-991A-30343	Sequence 30343, A
c 32	88.5	17.2	539	4	US-09-252-991A-21539	Sequence 21539, A
c 33	88.5	18.1	618	4	US-09-252-991A-23696	Sequence 23696, A
c 34	88	18.0	162	4	US-09-252-991A-19744	Sequence 19744, A
c 35	88	17.1	170	4	US-09-252-991A-22362	Sequence 22362, A
c 36	88	18.0	181	4	US-09-252-991A-18426	Sequence 18426, A
c 37	88	18.0	285	4	US-09-252-991A-20552	Sequence 20552, A
c 38	87.5	17.9	221	4	US-09-252-991A-27937	Sequence 27937, A
c 39	87.5	17.9	319	4	US-09-252-991A-21276	Sequence 21276, A
c 40	87.5	17.9	635	4	US-09-252-991A-32084	Sequence 32084, A
c 41	87	17.8	148	4	US-09-252-991A-23743	Sequence 23743, A
c 42	87	17.8	176	4	US-09-252-991A-29512	Sequence 29512, A
c 43	87	17.8	242	4	US-09-252-991A-18664	Sequence 18664, A
c 44	87	17.8	479	4	US-09-252-991A-32994	Sequence 32994, A
c 45	87	17.8	1476	4	US-09-252-991A-29427	Sequence 29427, A

ALIGNMENTS

RESULT 1
US-09-252-991A-32101
; Sequence 32101, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32101
; LENGTH: 162
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32101

Alignment Scores:
Pred. No.: 0.00488 Length: 162
Score: 101.00 Matches: 38
Percent Similarity: 32.85% Conservative: 7
Best Local Similarity: 27.74% Mismatches: 38
Query Match: 20.70% Indels: 54
DB: 4 Gaps: 6

US-10-042-417A-29 (1-278) x US-09-252-991A-32101 (1-162)

Qy 278 AATGCCCHCCAGTGGCGGCGCT----- 255

Db 1 SerAlaProProGlyAlaProSerArgProAlaAsnAlaGlySerGlyAlaAlaAala 20

QY 254 -----CCGCCAGCGCTGCGGAGATCCAGGTACGCTCCGATGGGTCCGCA 210
Db 21 AspArgArgAlaGlyProAlaHisArgArgArgGlyArgArgSerGlyAlaThr 40
QY 209 ATACTCTCGCA----- 198
Db 41 IleArgCysAlaArgSerArgSerValArgLysCysArgArgSerProProArgAlaPro 60
QY 197 -----CACACTCCCTCCATAAGCGGCACACGCGCCACCGCAGCAGC--- 153
Db 61 ArgThrThrProHisAlaProCysArgAlaGlySerAlaAlaProProAlaAlaAla 80
QY 152 -----CCTTGGCGGCAGAGGTGAGCACACGCTCCACCA----- 117
Db 81 HisArgSerAlaValGlyProGlyArgSerAlaArgProValArgGlyArgProSerGly 100
QY 116 CCTCGCCAGGT-----TACTCACACAGAGTNTCCCGGTCTACCGA 72
Db 101 SerProValGlyArgArgSerAlaArgSerAlaArgArgPro-----Gly 117
QY 71 GGANCCCGGACTCGCCGCAAGCANCCTACCGCTCCCTCACCAGCCC 21
Db 118 GlyHisArgAlaThrProArgCysAlaAlaLysSerAlaProAlaLeuPro 134
RESULT 2
US-09-252-991A-26586
; Sequence 26586, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26586
; LENGTH: 328
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26586
Alignment Scores:
Pred. No.: 0.00619 Length: 328
Score: 100.50 Matches: 31
Percent Similarity: 45.00% Conservative: 14
Best Local Similarity: 31.00% Mismatches: 26
Query Match: 19.51% Indels: 29
DB: 4 Gaps: 6
US-10-042-417A-29 (1-278) x US-09-252-991A-26586 (1-328)
QY 33 AATGGACCGCGTAGTCTGCGCGAGTCCCGGNTCCCGTAGACCGC----- 83
Db 132 AspGlyAlaAlaSerCysMetArgLeuLeuProGlyAsnThrLeuAspProProLeuPro 151
QY 84 -----CGGANACTTCGTGTGAGTAACCTGCGGAGGTGGTGAGCGGTGTCTACCT 137
Db 152 ArgSerLeuAlaProAlaCysArgValSerCysArg-----CysLeuPro 166
QY 138 TCC-----TGCCCGCCCAAGCGGT 155
Db 167 SerGlnAlaGluArgAspLysThrAlaArgArgGlyMetAspIleLeuProGlyArgGln 186
QY 156 TGCTGCGGTTGCGTGTGCGGTGTATGGAGGG---AGTGTGTGCGCAGAGTATTGC 212
Db 187 AspSerGlyArgProAlaAlaSerHisGlySerProAlaProCysLeuArgProCys 206
QY 213 GGACCCATCGGACGGTAACTGGA-----TCTCCGACGCGCTGGCGGAGCGCGCC 263

Db 207 ArgProThrSerAlaArgAlaAlaCysArgProSerProArg---TrpTyrArgProAla 225
RESULT 3
US-09-252-991A-16884
; Sequence 16884, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 16884
; LENGTH: 258
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-16884
Alignment Scores:
Pred. No.: 0.0125 Length: 258
Score: 97.50 Matches: 41
Percent Similarity: 42.48% Conservative: 7
Best Local Similarity: 36.28% Mismatches: 31
Query Match: 19.98% Indels: 34
DB: 4 Gaps: 8
US-10-042-417A-29 (1-278) x US-09-252-991A-16884 (1-258)
QY 278 AATGCCCCNCCAGTGGCCGG-----CCTCCG-----CCAGGC 246
Db 92 AspValProProGlyGlyArgGluSerProProArgSerGlySerProArgArgProGly 111
QY 245 -----CTCGGAGATCCAGGTACCTCCGATGGGTCCGCATA 207
Db 112 ProProHisAlaProLeuAlaAlaArgSerValArgArgSerAlaHisSerSerPro 131
QY 206 CTCCTGC---GCACACACTCTCCATTAAGCGGCACACGCGCCACCGCAGCAACGCT 150
Db 132 CysCysProAlaValPheProAlaValProAla-----ArgProValProGlyThrPro 149
QY 149 TGGCGGCGCAGGAAGGTGAGCA-----CAGGCTCCACCA 117
Db 150 TyrArgValArgProAlaAlaGlyProValProValGluAlaThrArgArgSerProPro 169
QY 116 CCT---CCGCCAGGTACTCAACACGAGGTNTCCCGGGGTCTACGGA----- 72
Db 170 ProAlaProProAlaValArgSerArgArgArgValAlaGlyArgAlaGlyThr 189
QY 71 GGANCCCGGACTCGCGCAAGCANCCTACCGCTCCCA 35
Db 190 Gly-ProAlaThrArgAlaAlaMetAlaProAlaPro 201
RESULT 4
US-09-252-991A-32112
; Sequence 32112, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32112
; LENGTH: 258
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32112

272 CCNCGAGTGGCGGGCTCCGCCAGGCCTCGCGAGATCCAGGTACCTCGCA
 527 ProValaIaGlyGlnProaIaAlaGlyHisArgaArgIeuArgaArgPro
 212 GCANTACTCTCGGCACACTCCCTCCATAAGCGGCACACCGAGGCCACCGCGC

```
Db 546 -----GlnGlyArgGlnProAspProAlaLeuPro 555
Qy 152 CTTGGCGGCGAGAGGTGAGCACGCTCCACCA-----CCTCGCGCCAGGTTACTCA 99
Db 556 ProTyrArg-----ProProArgHisProAlaProGlyTyrArg 568
Qy 98 ACAGAGAGTNTCCGCGGGTC-----TACGGAGGANCOCGGGACTCGCGCAAGCA 48
Db 569 ArgGlnArgProAspArgLeuAlaValAlaValArgLeuArgProArgArgSerAla 588
Qy 47 NCTACCGCTCCATTCTCTACCACCGCGCC 18
Db 589 GluProGlyGlnHisProGlyGlnProAla 598

RESULT 10
US-09-252-991A-19409
; Sequence 19409, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19409
; LENGTH: 165
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19409

Alignment Scores:
Pred. No.: 0.0401 Length: 165
Score: 92.50 Matches: 34
Percent Similarity: 39.22% Conservative: 6
Best Local Similarity: 33.33% Mismatches: 24
Query Match: 18.95% Indels: 38
DB: 4 Gaps: 6

US-10-042-417A-29 (1-278) x US-09-252-991A-19409 (1-165)
Qy 269 CCAGTGGCGGCTCCGCCAGGCTCGGAGATCCAGGTACGCTCCGATGGTTCGCGCA 210
Db 10 ProGlySerArgProAlaAsp-----ArgAlaThrGly----- 20
Qy 209 ATACTCTGCGCACACACTCCCTCCATAAGCGGCACAGCGCCAGCCCGCAGCAACGCCT 150
Db 21 -----CysAlaHisSerProAlaGlyArgAlaArgArgProArg----- 34
Qy 149 TGGCGGCGAGAGGTGAGCACAGCTTCCA-----CCACCTCCGCCCA 108
Db 35 --ArgAlaAlaArgProGlyHisAlaProGlyArgHisThrGlyAlaProArgPro 53
Qy 107 GGTACTCAACAGCAAGTNTCCGCGGTCT----- 77
Db 54 SerArgAlaArgArgArgAlaPro-AlaSerAlaArgAlaTyrLeuArgGlyLeuProAl 73
Qy 76 ---ACGGAGGANCOCGGGACTCGCGCAAGCANCTACCGGCTCCATTCTCCACGAGCC 21
Db 73 aGlyThrArgThrProGlyAlaArg-----ProAlaProArgSerSerAlaPr 89
Qy 20 GCCG 17
Db 89 opro 90

RESULT 11
US-09-252-991A-22682
; Sequence 22682, Application US/09252991A
```

```
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22682
; LENGTH: 457
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22682

Alignment Scores:
Pred. No.: 0.0472 Length: 457
Score: 92.50 Matches: 24
Percent Similarity: 45.00% Conservative: 3
Best Local Similarity: 40.00% Mismatches: 22
Query Match: 18.95% Indels: 11
DB: 4 Gaps: 3

US-10-042-417A-29 (1-278) x US-09-252-991A-22682 (1-457)
Qy 275 GCCCNCACAGTGGCGGCTCCGCCA-----GGCCTGCGGAGATCCAGGTTAGCG 225
Db 16 AlaProProAlaAlaArgProProAlaGlyThrThrGlyTrpIleArg 35
Qy 224 TCCGATGG;TCCGCAATACTCTGCGCACACACTCCCTCCATAAGCGGCACACGAGGCCA 165
Db 36 AsnGlySerSerCysArgArgSerAlaHisSer-----ArgProAlaArgArgSer 52
Qy 164 CCCGCGAGCACGCTTGGCGGCGAGGAGGTGAGCACACGCTCCACACCTCCGCCAGGT 105
Db 53 ProAlaGlnProAlaTyrArgPro-----AlaProProProGluProGly 67

RESULT 12
US-09-252-991A-30868
; Sequence 30868, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30868
; LENGTH: 574
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30868

Alignment Scores:
Pred. No.: 0.049 Length: 574
Score: 92.50 Matches: 38
Percent Similarity: 35.38% Conservative: 8
Best Local Similarity: 29.23% Mismatches: 37
Query Match: 18.95% Indels: 47
DB: 4 Gaps: 6

US-10-042-417A-29 (1-278) x US-09-252-991A-30868 (1-574)
```


Search completed: August 28, 2003, 13:29:16
Job time : 23.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model;

Run on: August 28, 2003, 07:12:45 ; Search time 49.5 Seconds
(without alignments)
1782.868 Million cell updates/sec

Title: US-10-042-417A-29

Perfect score: 515

Sequence: 1 ccdtagtactgnttcggc.....cggccacctgngggcgatt 278

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 2215726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlh
-O=/cgn2_1/USPNC_spool/US10042417/runat_19082003_133535_9502/app_query.fasta_1.455
-DB=A_Geneseq_19Jun03 -QWMT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10042417 -CGN_1.1.64 @runat_19082003_133535_9502 -LCPU=3
-NO_MMAP -LARGESQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A_Geneseq_19Jun03.*

1: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-embl/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-embl/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-embl/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-embl/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	436.5	84.8	91	21	AA83078	F-box protein FBP-
2	436.5	84.8	91	23	AAO22464	Human F-box protei
3	355	68.9	404	22	AB11672	Human secreted pro
4	355	68.9	404	22	AA79708	Human protein SEQ
5	349	67.8	276	22	AA84869	Human protein sequ
6	349	67.8	403	22	AA78724	Human protein SEQ
7	349	67.8	403	23	AB77551	F-box protein 44.3
8	349	67.8	403	23	AB77231	Human human protei
9	349	67.8	403	24	ABP58344	Human cell growth,
10	229	44.5	44	21	AA83058	F-box motif of FBP
11	229	44.5	44	23	AAO22482	Human F-box motif
12	98	19.0	161	22	ABG04119	Novel human diagno
13	98	19.0	161	22	ABG04122	Novel human diagno
14	94.5	19.4	266	22	ABG18528	Novel human diagno
15	94.5	19.4	926	22	ABG65135	Drosophila melanog
16	93	18.1	99	22	AB17496	Human nervous syst
17	91	18.6	680	23	ABJ05575	Breast cancer-asso
18	91	18.6	680	24	ABR7414	Breast cancer-asso
19	91	18.6	680	24	ABU36618	Lung cancer-associ
20	90	18.4	1938	24	ABP76682	Streptomyces virid
21	89.5	18.3	130	22	AAO01273	Human polypeptide
22	89.5	18.3	1938	24	ABP76680	Streptomyces virid
23	89	18.2	138	22	ABG22949	Novel human diagno
24	89	18.2	317	22	AA40791	Human polypeptide
25	89	18.2	317	22	AA40792	Human polypeptide
26	89	18.2	1938	24	ABP76681	Streptomyces virid
27	88.5	18.1	125	22	AAO00511	Human polypeptide
28	88.5	18.1	247	24	ABU11891	Human ABCA1 intera
29	87.5	17.9	99	22	AAO22445	Human polypeptide
30	87.5	17.9	165	22	ABG04414	Novel human diagno
31	87.5	17.0	208	22	ABG19244	Novel human diagno
32	87.5	17.9	432	23	ABG95649	Human nucleic acid
33	87.5	17.9	515	23	ABG96378	Human ovarian canc
34	87	17.8	708	23	ABG91504	Herbicide-activ
35	86	17.6	132	22	AAO06372	Human polypeptide
36	85.5	17.5	193	22	ABG31899	Peptide #4550 enco
37	85.5	17.5	193	22	ABG37137	Peptide #4643 enco
38	85.5	17.5	193	22	ABG22450	Protein #4449 enco
39	85.5	17.5	193	22	AA57857	Human bone marrow
40	85.5	17.5	193	22	AA70275	Human brain expres
41	85.5	17.5	193	22	AA18102	Peptide #4536 enco
42	85.5	17.5	193	22	AA30614	Peptide #4651 enco
43	85.5	17.5	193	22	AA05737	Peptide #4419 enco
44	85.5	17.5	193	23	ABG39918	Human peptide enco
45	85.5	17.5	658	22	ABG27820	Novel human diagno

ALIGNMENTS

RESULT 1
AA83078

ID AA83078 standard; Protein; 91 AA.

AC AA83078;

DT 16-AUG-2000 (first entry)

DE F-box protein FBP-10.

KW F-box protein; FBP; diagnosis; treatment; screening; agonist;
KW antagonist; proliferative disorder; differentiative disorder;
KW breast cancer; prostate cancer; ovarian cancer; cancer;
KW small cell lung carcinoma; immune disorder; cardiovascular disorder;
KW inflammatory disorder; human.

XX Homo sapiens.

XX

EH Key Location/Qualifiers
 FT Misc-difference 15 /note= "Unidentified amino acid"
 FT Misc-difference 22 /note= "Unidentified amino acid"
 FT Misc-difference 28 /note= "Unidentified amino acid"
 FT Misc-difference 89 /note= "Unidentified amino acid"
 FT Misc-difference 89 /note= "Unidentified amino acid"

XX WO200012679-A1.

XX PN 09-MAR-2000.

XX PD 27-AUG-1999; 99WO-US19560.

XX PF 28-AUG-1998; 98US-0098355.

XX PR 03-FEB-1999; 99US-0118568.

XX PR 15-MAR-1999; 99US-0124449.

XX XX (UYNV) UNIV NEW YORK STATE.

XX PA Chlaure DS, Pagano M, Latres E;

XX PI WPI: 2000-256635/22.

XX DR N-PSDB; AA293360.

XX XX Novel nucleic acid for screening compounds useful for treating

PT proliferative and differentiative disorders such as cancer and immune

PT disorders comprises sequences encoding ubiquitin ligases

XX Claim 10; Figure 13a; 245pp; English.

XX CC Nucleic acids encoding substrate-targeting subunits of ubiquitin

CC ligases with F-box motifs (F-box proteins) are useful for diagnosis

CC of proliferative and differentiated related disorders by measuring

CC FBP gene expression. Cells expressing such proteins or

CC their fragments are useful for screening compounds. The compounds

CC are agonists or antagonists, which are useful for treating a

CC proliferative or differentiative disorder in a mammal such as

CC breast, ovarian and prostate cancer and small cell lung carcinoma

CC and also major opportunistic infections, immune disorders,

CC cardiovascular diseases and inflammatory disorders, FBP protein,

CC analogs, derivatives and their subsequences, anti-FBP antibodies

CC are also useful in diagnosis of the disorders.

XX SQ Sequence 91 AA;

Alignment Scores:

Pred. No.: 9,06e-40 Length: 91

Score: 436.50 Matches: 91

Percent Similarity: 98.91% Conservative: 0

Best Local Similarity: 98.91% Mismatches: 0

Query Match: 84.76% Indels: 1

DB: 21 Gaps: 1

US-10-042-417A-29 (1-278) x AA83078 (1-91)

QY 2 CGTAGTACTGNTTCGGCGCGCTGGTGAGGAATGGAGCGGTAGNTGCTTGGCGGAGT 61

DB 1 ArgSerThrGlyPheArgArgAlaGlyGluGluTrpSerArg----LeuAlaAlaSer 19

QY 62 CCCGGGNTCCCTCCAGACCGCGGAGACCTTCGTGTGAGTAACCTGGCGGAGGTGGT 121

DB 20 ProGly***LeuArgArgProAla***ThrPheValLeuSerAsnLeuAlaGluValVal 39

QY 122 GAGCGTGTGCTACCTTCCTCCCGCGCAAGCGCTGCTGCGGTGGCGCTGGCTGCCGC 181

DB 40 GluArgValLeuThrPheLeuProAlaAlaValLeuLeuArgValAlaCysValCysArg 59

QY 182 TTATGAGGAGGTGTGTGCGCAGAGTATTGCGGACCCATCGGACCGCTTAACCTGGATCTCC 241

DB 60 LeuTrpArgGluCysValArgValLeuArgValLeuArgThrHisArgSerValThrTrpIleSer 79

QY 242 GCAGCGCTGGCGGAGCGCGCGCACCTGGNGGGCAT 277
 DB 80 AlaGlyLeuAlaGluAlaGlyHisLeu***GlyHis 91

RESULT 2

AAO22464

ID AAO22464 standard; Protein; 91 AA.

XX AAO22464;

XX 11-OCT-2002 (first entry)

XX Human F-box protein FBPI0 SEQ ID No 30.

XX Cytostatic; immunomodulator; cardiant; antiinflammatory; antimicrobial;

XX proliferative; differentiative disorder; Skp2; F-box protein; cancer;

XX ubiquitin ligase; breast cancer; prostate cancer; ovarian cancer;

XX small cell lung carcinoma; immune disorder; parathyroid adenoma; FBP;

XX inflammatory disorder; lymphoma; major opportunistic infection;

XX certain cardiovascular disease; human.

XX Homo sapiens.

XX WO200255665-A2.

XX 18-JUL-2002.

XX 07-JAN-2002; 2002WO-US00311.

XX 05-JAN-2001; 2001US-260179P.

XX (UYNV) UNIV NEW YORK STATE.

XX Pagano M;

XX WPI: 2002-599665/64.

XX N-PSDB; AAL41051.

XX Screening compounds for treating proliferative disorders, e.g. breast

XX cancer or prostate cancer, infections or immune disorders, comprises

XX detecting a change in the activity of Skp2 with either p27 or Cks1

XX Disclosure; Fig 13; 246pp; English.

XX The invention relates to screening compounds useful for the treatment of

XX proliferative or differentiative disorders comprising detecting a change

XX in the activity of Skp2 (F-box protein). The method is useful for

XX screening compounds for the treatment of proliferative or differentiative

XX disorders, particularly cancer. These compounds include small molecules,

XX or compounds or derivatives or analogues of the new ubiquitin ligases.

XX The compounds are useful for treating diseases such as cancer (e.g.

XX breast cancer, prostate cancer or ovarian cancer, lymphoma, small cell

XX lung carcinoma or parathyroid adenomas), major opportunistic infections,

XX immune disorders, certain cardiovascular diseases or inflammatory

XX disorders. This sequence represents an F-box protein (FBP) relating to

XX the invention.

XX SQ Sequence 91 AA;

Alignment Scores:

Pred. No.: 9,06e-40 Length: 91

Score: 436.50 Matches: 91

Percent Similarity: 98.91% Conservative: 0

Best Local Similarity: 98.91% Mismatches: 0

Query Match: 84.76% Indels: 1

DB: 23 Gaps: 1

US-10-042-417A-29 (1-278) x AAO22464 (1-91)

QY 2 CGTAGTACTGNTTCGGCGCGCTGGTGAGGAATGGAGCGGTAGNTGCTTGGCGGAGT 61

DB 1 ArgSerThrGlyPheArgArgAlaGlyGluGluTrpSerArg----LeuAlaAlaSer 19

Alignment Scores:
 Pred. No.: 4.57e-30 Length: 276
 Score: 349.00 Matches: 76
 Percent Similarity: 92.68% Conservative: 0
 Best Local Similarity: 92.68% Mismatches: 6
 Query Match: 67.77% Indels: 2
 DB: 22 Gaps: 0

US-10-042-417A-29 (1-278) x AAB94869 (1-276)

Qy 34 ATGAGCCGCTAGTCCGCGGAGTCCCG-GGNTCTCCGTAGACCGCGGAN-ACC 91
 Db 1 MetGluProValGlyCysGlyGluCysArgGlySerValAspProArgSerThr 20
 Qy 92 TTCGTGTGAGTAACTGCGGAGGTGCTGAGGCTGCTCCTCTCCGCGGCAAG 151
 Db 21 PheValLeuSerAsnLeuAlaGluValGluValGluValPheLeuProAlaLys 40
 Qy 152 GCGTTGCTGCGGCTGCGTGTGCGCTTATGAGGAGGTGCTGCGCAGAGTATTG 211
 Db 41 AlaLeuLeuArgValAlaCysValCysArgLeuTrpArgGluCysValArgValLeu 60
 Qy 212 CGGACCATCGGACCGTAACCTGATCTCCGAGGCTGCGGAGGCGCGCCACCTGGNG 271
 Db 61 ArgThrHisArgSerValThrTrpPileSerAlaGlyLeuAlaGluAlaGlyHisLeuGlu 80
 Qy 272 GGGCAT 277
 Db 81 GlyHis 82

RESULT 6
 AAM78724
 ID AAM78724 standard; Protein; 403 AA.
 XX
 AC AAM78724;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human protein SEQ ID NO 1386.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation.
 XX
 OS Homo sapiens.
 XX
 PN WO200157190-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 05-FEB-2001; 2001WO-US04098.
 XX
 PR 03-FEB-2000; 2000US-0496914.
 PR 27-APR-2000; 2000US-0560875.
 PR 20-JUN-2000; 2000US-0598075.
 PR 19-JUL-2000; 2000US-0620325.
 PR 01-SEP-2000; 2000US-0654936.
 PR 15-SEP-2000; 2000US-0663561.
 PR 20-OCT-2000; 2000US-0693325.
 PR 30-NOV-2000; 2000US-0728422.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
 PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
 XX
 DR WPI: 2001-476283/51.
 DR N-PSDB; AAK51857.
 XX
 PT Nucleic acids encoding polypeptides with cytokine-like activities,

PT useful in diagnosis and gene therapy -
 XX
 PS Claim 20; Page 3(47-3648; 6221pp; English.
 XX
 CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity relating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
 CC (AAM80020) are omitted as the relevant pages from the sequence listing
 CC were missing at the time of publication.
 XX

Sequence 403 AA;
 Alignment Scores:
 Pred. No.: 4.72e-30 Length: 403
 Score: 349.00 Matches: 76
 Percent Similarity: 92.68% Conservative: 0
 Best Local Similarity: 92.68% Mismatches: 6
 Query Match: 67.77% Indels: 2
 DB: 22 Gaps: 0

US-10-042-417A-29 (1-278) x AAM78724 (1-403)

Qy 34 ATGAGCCGCTAGTCCGCGGAGTCCCG-GGNTCTCCGTAGACCGCGGAN-ACC 91
 Db 1 MetGluProValGlyCysGlyGluCysArgGlySerValAspProArgSerThr 20
 Qy 92 TTCGTGTGAGTAACTGCGGAGGTGCTGAGGCTGCTCCTCTCCGCGGCAAG 151
 Db 21 PheValLeuSerAsnLeuAlaGluValGluValGluValPheLeuProAlaLys 40
 Qy 152 GCGTTGCTGCGGCTGCGTGTGCGCTTATGAGGAGGTGCTGCGCAGAGTATTG 211
 Db 41 AlaLeuLeuArgValAlaCysValCysArgLeuTrpArgGluCysValArgValLeu 60
 Qy 212 CGGACCATCGGACCGTAACCTGATCTCCGAGGCTGCGGAGGCGCGCCACCTGGNG 271
 Db 61 ArgThrHisArgSerValThrTrpPileSerAlaGlyLeuAlaGluAlaGlyHisLeuGlu 80
 Qy 272 GGGCAT 277
 Db 81 GlyHis 82

RESULT 7
 ABB77551
 ID ABB77551 standard; Protein; 403 AA.
 XX
 AC ABB77551;
 XX
 DT 08-AUG-2002 (first entry)
 XX
 DE F-box protein 44.33 SEQ ID NO 2.
 XX
 KW F-box protein; cytostatic; virucidal; immunomodulatory; HIV;
 KW antiinflammatory; haemostatic; malignant tumour; infection;
 KW human immunodeficiency virus; immunological disease; gene therapy.
 XX
 OS Unidentified.
 XX
 PN WO200232951-A1.
 XX
 PD 25-APR-2002.
 XX
 PF 21-SEP-2001; 2001WO-CN01443.
 XX

PR 22-SEP-2000; 2000CN-0125325.
 PA (BIOW-) BLOWDOWN GENE DEV INC SHANGHAI.
 PI Mao Y, Xie Y;
 XX
 XX
 XX WPI; 2002-340233/37.
 DR N-PSDB; ABL99954.
 XX
 XX
 PT A novel polypeptide, a F-box protein 44.33 and encoding polynucleotide,
 PT used in diagnosis and treatment of diabetes, menstrual disturbance,
 PT peptic ulcer, arrhythmia, hemophthysis and epilepsy -
 XX
 XX Claim 1; Page 28; 36pp; Chinese.
 XX
 CC The invention relates to F-box protein 44.33 with cytostatic,
 CC virucidal, immunomodulatory, antiinflammatory and haemostatic
 CC activity. The protein and encoding polynucleotide are used in diagnosis
 CC and treatment of malignant tumour, haemopathy, human immunodeficiency
 CC virus (HIV) infection, immunological diseases and various inflammations.
 CC The polynucleotide is useful in gene therapy.
 XX
 SQ Sequence 403 AA;
 Alignment Scores:
 Pred. No.: 4.72e-30 Length: 403
 Score: 349.00 Matches: 76
 Percent Similarity: 92.68% Conservative: 0
 Best Local Similarity: 92.68% Mismatches: 6
 Query Match: 67.77% Indels: 2
 DB: 23 Gaps: 0
 US-10-042-417A-29 (1-278) x ABB77551 (1-403)
 QY 34 ATGGAGCCGAGTGTCTGCGCGAGTCCCG-GGNTCTCTCGTAGACCCGCGGAN-ACC 91
 Db 1 MetGluProValGlyCysCysGlyGluCysArgGlySerValAspProArgSerThr 20
 QY 92 TTCGTGTTCAGTAACCTGCGCGAGTGTGGAGCGTGTGCTCACCTTCTGCCCCCAAG 151
 Db 21 PheValLeuSerAsnLeuAlaGluValValGluArgValLeuThrPheLeuProAlaLys 40
 QY 152 GCGTTGCTGCGGGTGGCTGTGCTGCTTATGGAGGAGTGTGCGCGAGAGATTG 211
 Db 41 AlaLeuLeuArgValAlaCysValCysArgLeuTrpArgGluCysValArgValLeu 60
 QY 212 CGGACCCATCGAGCGTAACCTGCGAGTGTCTGCGAGCGCTGCGGAGCGCCACCTGGNG 271
 Db 61 ArgThrHisArgSerValThrTrpIleSerAlaGlyLeuAlaGluAlaGlyHisLeuGlu 80
 QY 272 GGGCAT 277
 Db 81 GlyHis 82
 RESULT 8
 ABB97231
 ID ABB97231 standard; Protein; 403 AA.
 AC ABB97231;
 XX
 XX 27-JUN-2002 (first entry)
 XX
 XX Novel human protein SEQ ID NO: 499.
 XX
 XX Human; antianaemic; vulnary; antiinflammatory; immunomodulator;
 KW antifertility; cerebroprotective; cytostatic; rheumatic; gene therapy;
 KW neuroprotective; antiparkinsonian; protein therapy; EST;
 KW expressed sequence tag.
 XX
 OS Homo sapiens.
 XX
 XX WO200222660-A2.
 PN
 XX

21-MAR-2002.
 PD XX
 PF 10-SEP-2001; 2001WO-US26015.
 XX
 PR 11-SEP-2000; 2000US-0659671.
 XX
 XX (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
 PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;
 XX
 XX WPI; 2002-292408/33.
 DR N-PSDB; ABN32417.
 XX
 XX An isolated polynucleotide for treating diseases associated with its
 PT encoded polypeptide such as cancer and multiple sclerosis -
 XX
 PS Example 2; SEQ ID NO 499; 509pp; English.
 XX
 CC The present invention provides the protein and coding sequences of 444
 CC novel human proteins. These were isolated from expressed sequences tags
 CC (ESTs). They can be used to stimulate cell growth, to regulate
 CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
 CC e.g. in burn treatment, to regulate the immune system e.g. to treat
 CC multiple sclerosis, to regulate activin or inhibin e.g. to treat
 CC infertility, to regulate haemostasis or thrombolysis e.g. to treat
 CC stroke and cancer, to screen for drugs, to treat inflammatory conditions
 CC e.g. rheumatoid arthritis, and to treat nervous system disorders e.g.
 CC Parkinson's disease. The present sequence is a protein of the invention.
 XX
 SQ Sequence 403 AA;
 Alignment Scores:
 Pred. No.: 4.72e-30 Length: 403
 Score: 349.00 Matches: 76
 Percent Similarity: 92.68% Conservative: 0
 Best Local Similarity: 92.68% Mismatches: 6
 Query Match: 67.77% Indels: 2
 DB: 23 Gaps: 0
 US-10-042-417A-29 (1-278) x ABB97231 (1-403)
 QY 34 ATGGAGCCGAGTGTCTGCGCGAGTCCCG-GGNTCTCTCGTAGACCCGCGGAN-ACC 91
 Db 1 MetGluProValGlyCysCysGlyGluCysArgGlySerValAspProArgSerThr 20
 QY 92 TTCGTGTTCAGTAACCTGCGCGAGTGTGGAGCGTGTGCTCACCTTCTGCCCCCAAG 151
 Db 21 PheValLeuSerAsnLeuAlaGluValValGluArgValLeuThrPheLeuProAlaLys 40
 QY 152 GCGTTGCTGCGGGTGGCTGTGCTGCTTATGGAGGAGTGTGCGCGAGAGATTG 211
 Db 41 AlaLeuLeuArgValAlaCysValCysArgLeuTrpArgGluCysValArgValLeu 60
 QY 212 CGGACCCATCGAGCGTAACCTGCGAGTGTCTGCGAGCGCTGCGGAGCGCCACCTGGNG 271
 Db 61 ArgThrHisArgSerValThrTrpIleSerAlaGlyLeuAlaGluAlaGlyHisLeuGlu 80
 QY 272 GGGCAT 277
 Db 81 GlyHis 82
 RESULT 9
 ABB58344
 ID ABB58344 standard; Protein; 403 AA.
 XX
 XX ABB58344;
 XX
 XX 07-APR-2003 (first entry)
 XX
 XX Human cell growth, differentiation and death protein CGDD-15.
 DE
 XX CGDD-15; cell growth; cell differentiation; cell death; human;
 KW

KW cytostatic; antiarteriosclerotic; hepatotropic; antiinflammatory;
KW antipsoriatic; antineoplastic; ophthalmological; auditory;
KW anticonvulsant; cerebrotective; nootropic; neuroprotective;
KW antiparkinsonian; neuroleptic; tranquilizer; immunosuppressive;
KW anti-HIV; antiallergic; antisthmatic; antihypertensive; antidiabetic;
KW dermatological; nephrotoxic; antirheumatic; antarthritic;
KW antifungal; virucide; antibacterial; fungicide;
KW antiparasitic; protozoacide; antihelminthic; antiinfertility;
KW gynaecological; gene therapy.
XX
OS Homo sapiens.
XX
XX WO200297032-A2.
XX
XX
XX 05-DEC-2002.
XX
XX 05-APR-2002; 2002WO-US11152.
XX
XX 06-APR-2001; 2001US-282110P.
XX 11-APR-2001; 2001US-283294P.
XX 26-APR-2001; 2001US-286820P.
XX 27-APR-2001; 2001US-287228P.
XX 16-MAY-2001; 2001US-291662P.
XX 18-MAY-2001; 2001US-291846P.
XX 25-MAY-2001; 2001US-293727P.
XX 01-JUN-2001; 2001US-295263P.
XX 01-JUN-2001; 2001US-295340P.
XX 15-JAN-2002; 2002US-349705P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Azimzal Y, Au-Young JK, Batra S, Baughn MR, Becha SD, Borowsky ML;
XX Burford N, Ding L, Elliott VS, Emerling BM, Gandhi AR, Gietzen KJ;
XX Griffin JA, Hafalla AJA, Honchell CD, Lal PG, Lee SY, Lu DAM;
XX Arvizu CS, Ramkumar J, Reddy R, Sanjanwala MM, Rang YT, Wallia NK;
XX Wang YE, Warren BA, Xu Y, Yang J, Yao MG, Yue H, Zebbarjadian Y;
XX WPI; 2003-140453/13.
XX
XX N-PSDB; ABZ24703.
XX
XX Novel human proteins associated with cell growth, differentiation and
XX death, useful for treating, diagnosing or preventing cancer,
XX developmental, neurological, reproductive or autoimmune/inflammatory
XX disorders -
XX
XX Claim 1; Page 203-204; 238pp; English.
XX
XX The present sequence is the protein sequence of human CGDD-15, a
XX novel protein associated with cell growth, differentiation and
XX death. The sequence is predicted from Incyte clone 5565648CB1,
XX which was isolated from a foetal liver tissue cDNA library.
XX Structural features establish the protein as being associated with
XX cell growth, differentiation and death. The invention is based on
XX novel human CGDD-1 to -21 proteins (see ABP58330-50), the
XX polynucleotides encoding them (see ABZ24689-709), and to the use of
XX these for the diagnosis, treatment or prevention of cell
XX proliferative disorders, including cancer, developmental disorders,
XX neurological disorders, autoimmune disorders, reproductive
XX disorders, and disorders of the placenta, and in the assessment of
XX the effects of exogenous compounds on the activity and expression
XX of proteins and nucleic acids associated with cell growth,
XX differentiation and death.
XX
XX Sequence 403 AA;
XX
XX Alignment Scores:
XX Pred. No.: 4,72e-30 Length: 403
XX Score: 349.00 Matches: 76
XX Percent Similarity: 92.68% Conservative: 0
XX Best Local Similarity: 92.68% Mismatches: 6
XX Query Match: 67.77% Indels: 2
XX DB: 24 Gaps: 0

US-10-042-417a-29 (1-278) x ABP58344 (1-403)
QY 34 ATGGAGCGGTAGTGTGGGGAGTCCG -GGTCTCCGTAGACCGCGGAN-ACC 91
Db 1 MetGluFroValGlyCysCysGlyGluCysArgGlySerSerValAspProArgSerThr 20
QY 92 TTCGTGTGTAGTAACCTGGGGAGGTGGTGGAGGCTGTCTCACCCTTCTCCGCCGCCAAG 151
Db 21 PheValLeuSerAsnLeuAlaGluValValGluArgValLeuThrPheLeuProAlaLys 40
QY 152 CGCTTGTCTGGGGTGGCTGGCTGGCTTATGGAGGGAGTGTGGCGGAGATATTG 211
Db 41 AlaLeuLeuArgValAlaCysValCysArgLeuTrpArgGluCysValArgArgValLeu 60
QY 212 CGGACCATCGGAGCGTAACCTGGATCTCCGACGCTGGCGGAGGCGGCGACCTGGNG 271
Db 61 ArgThrHisArgSerValThrTrpIleSerAlaGlyLeuAlaGlyHisLeuGlu 80
QY 272 GGCAT 277
Db 81 GlyHis 32
RESULT 10
AAY83058
ID AAY83058 standard; Peptide; 44 AA.
XX
XX AAY83058;
XX
XX 16-AUG-2000 (first entry)
XX
XX F-box motif of FBP-10.
XX
XX F-box protein; FBP; diagnosis; treatment; screening; agonist;
XX antagonist; proliferative disorder; differentiative disorder;
XX breast cancer; prostate cancer; ovarian cancer; cancer;
XX small cell lung carcinoma; immune disorder; cardiovascular disorder;
XX inflammatory disorder; human.
XX
XX Homo sapiens.
XX
XX WO200012679-A1.
XX
XX 09-MAR-2000.
XX
XX 27-AUG-1999; 95WO-US19560.
XX
XX 28-AUG-1998; 95US-0098355.
XX 03-FEB-1999; 95US-0118568.
XX 15-MAR-1999; 95US-0124449.
XX
XX (UYNV) UNIV NEW YORK STATE.
XX
XX Chaur DS, Pagano M, Latres E;
XX WPI; 2000-256635/22.
XX
XX Novel nucleic acid for screening compounds useful for treating
XX proliferative and differentiative disorders such as cancer and immune
XX disorders comprises sequences encoding ubiquitin ligases -
XX
XX Disclosure: Page 232; 245pp; English.
XX
XX Nucleic acids encoding substrate-targeting subunits of ubiquitin
XX ligases with F-box motifs (F-box proteins) are useful for diagnosis
XX of proliferative and differentiative related disorders by measuring
XX FBP gene expression. Cells expressing such proteins or
XX their fragments are useful for screening compounds. The compounds
XX are agonists or antagonists, which are useful for treating a
XX proliferative or differentiative disorder in a mammal such as
XX breast, ovarian and prostate cancer and small cell lung carcinoma
XX and also major opportunistic infections, immune disorders,
XX cardiovascular diseases and inflammatory disorders. FBP protein,
XX analogs, derivatives and their subsequences, anti-FBP antibodies

CC are also useful in diagnosis of the disorders.

XX Sequence 44 AA;

Alignment Scores:

Pred. No.: 6.86e-17 Length: 44
Score: 229.00 Matches: 44
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 44.47% Indels: 0
DB: 21 Gaps: 0

US-10-042-417A-29 (1-278) x AAY83058 (1-44)

QY 107 CTGCGGAGGTGGTGGAGCGTGTCTCACCTTCTCGCCGCAAGCGTGTCTCGGGTG 166
DB 1 LeuAlaGluValValGluArgValLeuThrPheLeuProAlaLysAlaLeuLeuArgVal 20
QY 167 GCCTGGGTGCGCGCTTATGAGGAGGTGTGTGCGCAGAGTATTGCGGACCCATCGGAGC 226
DB 21 AlaCysValCysArgLeuTrpArgGluCysValArgValLeuArgThrHisArgSer 40
QY 227 GTAACCTGGATC 238
DB 41 ValThrTrpIle 44

RESULT 11

AAO22482

ID AAO22482 standard; Protein; 44 AA.

XX AAO22482;

DT 11-OCT-2002 (first entry)

XX Human F-box motif amino residues of FBP10 SEQ ID No 63.

XX Cytostatic; immunomodulator; cardiant; antinflammatory; antimicrobial;
KW proliferative; differentiative disorder; Skp2; F-box protein; cancer;
KW ubiquitin ligase; breast cancer; prostate cancer; ovarian cancer;
KW small cell lung carcinoma; immune disorder; parathyroid adenoma; FBP;
KW inflammatory disorder; lymphoma; major opportunistic infection;
KW certain cardiovascular disease; human.

XX Homo sapiens.

XX WO200255665-A2.

XX 18-JUL-2002.

XX 07-JAN-2002; 2002WO-US00311.

XX 05-JAN-2001; 2001US-260179P.

XX (UYNV) UNIV NEW YORK STATE.

XX Pagano M;

XX WPI; 2002-599665/64.

XX Screening compounds for treating proliferative disorders, e.g. breast
PT cancer or prostate cancer, infections or immune disorders, comprises
PT detecting a change in the activity of Skp2 with either p27 or Cks1 -

XX Disclosure; Fig 1; 246pp; English.

XX The invention relates to screening compounds useful for the treatment of
CC proliferative or differentiative disorders comprising detecting a change
CC in the activity of Skp2 (F-box protein). The method is useful for
CC screening compounds for the treatment of proliferative or differentiative
CC disorders, particularly cancer. These compounds include small molecules,
CC or compounds or derivatives or analogues of the new ubiquitin ligases.
CC The compounds are useful for treating diseases such as cancer (e.g.
CC breast cancer, prostate cancer or ovarian cancer, lymphoma, small cell

CC lung carcinoma or parathyroid adenomas), major opportunistic infections,
CC immune disorders, certain cardiovascular diseases or inflammatory
CC disorders. This sequence represents an F-box protein (FBP) relating to
CC the invention.

XX SQ Sequence 44 AA;

Alignment Scores:

Pred. No.: 6.86e-17 Length: 44
Score: 229.00 Matches: 44
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 44.47% Indels: 0
DB: 23 Gaps: 0

US-10-042-417A-29 (1-278) x AAO22482 (1-44)

QY 107 CTGCGGAGGTGGTGGAGCGTGTCTCACCTTCTCGCCGCAAGCGTGTCTCGGGTG 166
DB 1 LeuAlaGluValValGluArgValLeuThrPheLeuProAlaLysAlaLeuLeuArgVal 20
QY 167 GCCTGGGTGCGCGCTTATGAGGAGGTGTGTGCGCAGAGTATTGCGGACCCATCGGAGC 226
DB 21 AlaCysValCysArgLeuTrpArgGluCysValArgValLeuArgThrHisArgSer 40
QY 227 GTAACCTGGATC 238
DB 41 ValThrTrpIle 44

RESULT 12

ABG04119

ID ABG04119 standard; Protein; 161 AA.

XX AC ABG04119;

XX 13-FEB-2002 (first entry)

XX Novel human diagnostic protein #4110.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

XX 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX N-PSDB; AAS68306.

XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity

XX Claim 20; SEQ ID No 34478; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and

CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques

to restore normal activity of (II) or to treat disease states involving (III). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic amino acid sequences of the invention.

Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published/pct/sequences.

XX PN WO200175067-A2.
 XX PD 11-OCT-2001.
 XX PF 30-MAR-2001; 2001WO-US08631.
 XX PR 31-MAR-2000; 2000US-0540217.
 XX PR 23-AUG-2000; 2000US-0649167.
 XX PA (HYSE-) HYSEQ INC.
 XX PI Drmanac RT, Liu C, Tang YT;
 XX DR WPI; 2001-639362/73.
 XX DR N-PSDB; AAS82715.
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity
 XX Claim 20; SEQ ID NO 48887; 103pp; English.
 PS The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: the sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX SQ Sequence 266 AA;
 Alignment Scores:
 Pred. No.: 0.0567 Length: 266
 Score: 94.50 Matches: 36
 Percent Similarity: 34.45% Conservative: 5
 Best Local Similarity: 30.25% Mismatches: 45
 Query Match: 19.36% Indels: 33
 DB: 22 Gaps: 6
 US-10-042-417A-29 (1-278) x ABG18528 (1-266)
 QY 275 GCCCNCAGGTGGCGCGCTCCGCCAGGCTCGGAGATCCAGTTACCTCCGATGGG 216
 Db 46 AlapProAlaGlyProAlaAlaProAlaAlaProAlaArgGlnProProArgProAla 65
 QY 215 TCGCAATACCTCTCGGCACAC-----ACTCCCTCCATAAGC----- 180
 Db 66 GlnProArgSerCysProArgArgValArgProThrProAlaTrpSerProProGlyArg 85
 QY 179 -----GGCACACGAGCCACCGCA-----GCAACGCGCTGG 147
 Db 86 ArgGlyProArgArgSerArgArgProProGlyGlyProGlyProAlaAlaProLeu 105
 QY 146 CGGCGAGGAGGTGACACACGCTCCACACCTCCGCCAGGTACTCAACACCA----- 93
 Db 106 ArgLeuSerArg---AlaHisSerProAlaProProGlySerProTyArgProHis 124

QY 92 -----AGTNTCCGCGGTCTACGGAGGANCCTCGG 63
 Db 125 GlyAlaGlyThrSerValAlaProTrpThrArgProProAlaAlaArgGlyThrGluAla 144
 QY 62 GAC-----TCGCCCAAGCANCCTACCGGTCTCCATTCCTCACCAGCCGCGCGA 15
 Db 145 AspProGlyArgCysProValSerAlaProGlyThrAlaProGlyGlnMetArgGly 163
 RESULT 15
 ABB65135
 ID ABB65135 standard; Protein; 926 AA.
 XX AC ABB65135;
 XX DT 26-MAR-2002 (first entry)
 XX DE Drosophila melanogaster polypeptide SEQ ID NO 22197.
 XX KW Drosophila; developmental biology; cell signalling; insecticide;
 XX KW pharmaceutical.
 XX OS Drosophila melanogaster.
 XX PN WO200171042-A2.
 XX PD 27-SEP-2001.
 XX PF 23-MAR-2001; 2001WO-US09231.
 XX PR 23-MAR-2000; 2000US-191637P.
 XX PR 11-JUL-2000; 2000US-0614150.
 XX PA (PEKE) PE CORP NY.
 XX PI Venter JC, Adams M, Li PWD, Myers EW;
 XX DR WPI; 2001-656860/75.
 XX DR N-PSDB; ABL09238.
 XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions
 XX PS Disclosure; SEQ ID NO 22197; 21pp + Sequence Listing; English.
 XX CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABB57737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX SQ Sequence 926 AA;
 Alignment Scores:
 Pred. No.: 0.0634 Length: 926
 Score: 94.50 Matches: 37
 Percent Similarity: 40.83% Conservative: 12
 Best Local Similarity: 30.83% Mismatches: 37
 Query Match: 19.36% Indels: 35
 DB: 22 Gaps: 6
 US-10-042-417A-29 (1-278) x ABB65135 (1-926)
 QY 272 CCNCCAGGTGGC-----CGGCTCCGCCAGGCTCGGAGATCCAGTTACCTCCGAT 219
 Db 566 ProThrGlyGlyGlnGlnGlnProProGlyProGlnSerGln-TyrGlyPro 585

Search completed: August 28, 2003, 13:24:26
Job time : 53.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2003, 01:18:55 ; Search time 5392 Seconds
(without alignments)
1253.087 Million cell updates/sec

Title: US-10-042-417A-29

Perfect score: 278

Sequence: 1 ccgtagtactgnttcgcgc.....cggccacctgnggggcatt 278

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

- 1: em_estba:**
- 2: em_esthum:**
- 3: em_estin:**
- 4: em_estmu:**
- 5: em_estov:**
- 6: em_estpl:**
- 7: em_estro:**
- 8: em_htc:**
- 9: gb_est1:**
- 10: gb_est2:**
- 11: gb_htc:**
- 12: gb_est3:**
- 13: gb_est4:**
- 14: gb_est5:**
- 15: em_estfun:**
- 16: em_estom:**
- 17: em_gss_hum:**
- 18: em_gss_inv:**
- 19: em_gss_pln:**
- 20: em_gss_vrt:**
- 21: em_gss_fun:**
- 22: em_gss_mam:**
- 23: em_gss_mus:**
- 24: em_gss_pro:**
- 25: em_gss_rod:**
- 26: em_gss_phg:**
- 27: em_gss_vri:**
- 28: gb_gss1:**
- 29: gb_gss2:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	273	98.2	278	14	245775 HSCZTG121 n
2	244	87.8	904	9	AL559170 AL559170
3	242	87.1	1201	13	BX385474 BX385474
4	241.4	86.8	762	10	BG328756 BG328756

5	241.4	86.8	871	10	BF209864
6	237.4	85.4	975	9	AL519688
7	233.4	84.0	770	14	CB997897
8	233.4	84.0	899	14	CD110208
9	233.4	84.0	1201	13	EX401591
10	232	83.5	679	12	BI458126
11	232	83.5	736	12	BI918213
12	232	83.5	792	9	AU143434
13	232	83.5	801	10	BF691820
14	232	83.5	808	13	BU594927
15	232	83.5	819	10	BE889612
16	232	83.5	861	13	BU193184
17	232	83.5	905	12	BI458472
18	232	83.5	928	13	BU164363
19	232	83.5	1001	13	EX391620
20	231	83.1	699	9	AV704428
21	230.4	82.9	847	13	BU183077
22	230	82.7	551	10	BE613868
23	229	82.4	736	10	BE252051
24	228.8	82.3	1201	9	AL524130
25	228.4	82.2	852	10	BE744221
26	228.8	81.2	321	12	BM193850
27	222	79.9	908	12	BM449628
28	221.2	79.6	585	10	EG699818
29	221.2	79.6	636	12	BI553087
30	221	79.5	531	9	AL705237
31	220	79.1	845	14	CB960550
32	218	78.4	645	10	BE253270
33	217	78.1	471	9	AN163504
34	216	77.7	760	10	BE739043
35	215	77.3	668	10	BF210147
36	212	76.3	889	13	BU161306
37	212	76.3	893	13	BU175330
38	210.8	75.8	658	10	BE258718
39	210.4	75.7	904	13	BQ425363
40	208.8	75.1	1201	13	EX424296
41	206.8	74.4	1201	13	EX443770
42	206.4	74.2	886	10	BF211635
43	202	72.7	914	12	BI769404
44	200.4	72.1	996	13	BQ942618
45	199	71.6	876	13	BU160419

ALIGNMENTS

RESULT 1:

245775
LOCUS HSCZTG121 278 bp mRNA linear EST 14-NOV-1994
DEFINITION HSCZTG121 normalized infant brain cDNA Homo sapiens cDNA clone
c-ztg12, rRNA sequence.

ACCESSION 245775

VERSION 245775.1 GI:575009

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 278)

AUTHORS M.D., Duprat,S., Houlgatte,R., Jumeau,M.N., Lamy,B., Lorenzo,F.,
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 278)

Aufrey,C., Behar,G., Bois,F., Bouchier,C., da Silva,C., Devignes

Mitchell,H., Mariage-Samson,R., Pietu,G., Pouliot,Y.,

Sebastiani-Kabaktchis,C. and Tessier,A.

IMAGE: molecular integration of the analysis of the human genome

and its expression

C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)

JOURNAL 95277534

MEDLINE 7757816

PUBMED

COMMENT Contact: Genethon

Genexpress-Genethon

Genethon Centre de recherche sur le Genome Humain

1,rue de l'Internationale, BP60 91002 EVRY Cedex, FRANCE

Tel: 33169477800

Fax: 33160778698
Email: genexpres@genethon.fr
Single read.
Genexpres_library_id: C; Genexpres_sequence_id: ylc-ztg12
Seq primer: (-21)M13-universal.
Location/Qualifiers
1..278
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="c-ztg12"
/sex="Female"
/tissue_type="total brain"
/dev_stage="3 months old"
/clone_lib="normalized infant brain cDNA"
/note="Organ: brain; Vector: lafmid BA; Site_1: HindIII; Site_2: NotI; sex:Female; dev_stage=3 months old; isolate-muscular atrophy patient; tissue_type=total brain; total mRNA was oligo-(dT) primed and directionally cloned 5' -> 3' into the HindIII -> NotI sites of the lafmid BA vector. Clone library from B Soares, Psychiatry Dept. Columbia University, USA. Normalization_method: Bencio Soares, P.N.A.S in press"
BASE COUNT 36 a 74 c 107 g 56 t 5 others
ORIGIN

Query Match 98.2%; Score 273; DB 14; Length 278;
Best Local Similarity 100.0%; Pred. No. 5.2e-54;
Matches 278; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCGTAGTACTGNTTCCGGGGGCTGGTGGAGGAATGGAGCGGTAGNTGCTTGGCGGAG 60
Db 1 CCGTAGTACTGNTTCCGGGGGCTGGTGGAGGAATGGAGCGGTAGNTGCTTGGCGGAG 60
QY 61 TCCCGGNTCTCCGTAGACCCCGGAGACCTTCGTGTGAGTAACCTGCGGAGGTGGT 120
Db 61 TCCCGGNTCTCCGTAGACCCCGGAGACCTTCGTGTGAGTAACCTGCGGAGGTGGT 120
QY 121 GGAGCGTGTGCTACCTTCTGCGCCGCAAGCGCTTCTGCGGGTGGCCCTGCGTGTGCCG 180
Db 121 GGAGCGTGTGCTACCTTCTGCGCCGCAAGCGCTTCTGCGGGTGGCCCTGCGTGTGCCG 180
QY 181 CTTATGAGGAGGTGTGCGCAGAGTATTGCGGACCCATCGGAGCGTAACCTGGATCTC 240
Db 181 CTTATGAGGAGGTGTGCGCAGAGTATTGCGGACCCATCGGAGCGTAACCTGGATCTC 240
QY 241 CGCAGGCGTGGCGAGCGCGCCACCTGNGGGGCATT 278
Db 241 CGCAGGCGTGGCGAGCGCGCCACCTGNGGGGCATT 278

RESULT 2
AL559170
LOCUS
DEFINITION
AL559170 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
Homo sapiens cDNA clone CS0DJ014YJ01 5-PRIME, mRNA sequence.
ACCESSION
AL559170
VERSION
AL559170.2 GI:31283303
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE
Full-length cDNA libraries and normalization
JOURNAL
Unpublished
On Feb 15, 2001 this sequence version replaced gi:12904405.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
was normalized. Library was constructed by Life Technologies, a

division of Invitrogen. This sequence belongs to sequence cluster 7864.f For more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DJ014CE010P1&cluster=7864.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DJ014CE010P1.
Location/Qualifiers
1..904
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DJ014YJ01"
/cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
/clone_lib="JURKAT"
/clone_line="Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 220 a 198 c 258 g 224 t 4 others
ORIGIN

Query Match 87.8%; Score 244; DB 9; Length 904;
Best Local Similarity 93.6%; Pred. No. 3.7e-47;
Matches 250; Conservative 1; Mismatches 16; Indels 0; Gaps 0;
QY 12 GNTTCCGGGGCTGGTGGAGGAATGGAGCGGTAGNTGCTTGGCGGAGTCCCGGNTCC 71
Db 118 GGTTCGGCGGCTGGTGGAGGAATGGAGCGGTAGTGCCTCGCGCGAGTCCCGGCTCC 177
QY 72 TCCGTAGACCCCGGAGNACCTTCGTGTGAGTAACCTGCGGGTGGCGGAGTGGTGGAGCGTGC 131
Db 178 TCCGTAGACCCCGGAGNACCTTCGTGTGAGTAACCTGCGGGTGGCGGAGTGGTGGAGCGTGC 237
QY 132 TCACCTTCTCTGCGCCGCAAGCGCTTCTGCGGGTGGCGCTGCGCTTATGGAGGG 191
Db 238 TCACCTTCTCTGCGCCGCAAGCGCTTCTGCGGGTGGCGCTGCGCTTATGGAGGG 297
QY 192 AGTGTGTGCGCAGATATTGCGGACCCATCGGAGCGTAACCTGCGAGGCGCTGG 251
Db 298 AGTGTGTGCGCAGATATTGCGGACCCATCGGAGCGTAACCTGCGAGGATVTCGCGAGCGCTGG 357
QY 252 CGGAGCGCGCCACCTGNGGGGCATT 278
Db 358 CGGAGCGCGCCACCTGNGGGGCATT 384

RESULT 3
BX385474
LOCUS
DEFINITION
BX385474 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
Homo sapiens cDNA clone CS0DL011P23 5-PRIME, mRNA sequence.
ACCESSION
BX385474
VERSION
BX385474.1 GI:30460505
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE
Full-length cDNA libraries and normalization
JOURNAL
Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 7864.f For
more information about this cluster, see
http://www.genoscope.cns.fr/

cgi-bin/cluster.cgi?seq=CS0DL011CH12QPI&cluster=7864.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DL011CH12QPI.

FEATURES

Location/Qualifiers
1. 1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DL011VP23"
/cell_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 309 a 237 c 316 g 288 t 51 others
ORIGIN

Query Match 87.1%; Score 242; DB 13; Length 1201;
Best Local Similarity 95.5%; Pred. No. 1.1e-46;
Matches 256; Conservative 0; Mismatches 11; Indels 1; Gaps 1;
QY 12 GNTTCGGCGGGCTGGTGGAGAAATGGAGCCGGTAGNTGCTTGGCGGAGTCCCG-GGNTC 70
DB 83 GGTTCGGCGGGCTGGTGGAGAAATGGAGCCGGTAGGCTGCTGGCGGAGTGGCGGCGCTC 142
QY 71 CTCGGTAGACCCGGGAGACCTTCGTTGAGTAACCTGGCGGAGTGGTGGAGCGGTG 130
DB 143 CTCGGTAGACCCGGGAGACCTTCGTTGAGTAACCTGGCGGAGTGGTGGAGCGGTG 202
QY 131 CTCACCTTCCTGCCCGCCAAAGCGTTCGTCGGGTGGCGTGCCTGCTTATGGAGG 190
DB 203 CTCACCTTCCTGCCCGCCAAAGCGTTCGTCGGGTGGCGTGCCTGCTTATGGAGG 262
QY 191 GAGTGTGGCGAGAGTATTGGGACCCATCGGAGCGTAACCTGGATCTCCGAGGCGTG 250
DB 263 GAGTGTGGCGAGAGTATTGGGACCCATCGGAGCGTAACCTGGATCTCCGAGGCGTG 322
QY 251 GCGGAGGCGCGCCACCTGGNGGGGCATT 278
DB 323 GCGGAGGCGCGCCACCTGGAGGGGCATT 350

RESULT 4
BG328756 762 bp mRNA linear EST 27-FEB-2001
LOCUS 602427037F1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:4546640 5',
DEFINITION mRNA sequence.

ACCESSION BG328756
VERSION BG328756.1 GI:13135194
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 762)
TITLE NTH-MGC http://mgc.nci.nih.gov/

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov

Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: NIH Intramural Sequencing Center
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1233 row: 1 column: 09
High quality sequence stop: 753.

FEATURES

source
Location/Qualifiers
1. 762
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4546640"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_15"
/note="Organ: colon; Vector: pOTB7; Site: 1; XhoI; Site_2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)"
BASE COUNT 176 a 180 c 229 g 177 t
ORIGIN

Query Match 86.8%; Score 241.4; DB 10; Length 762;
Best Local Similarity 95.5%; Pred. No. 1.4e-46;
Matches 256; Conservative 0; Mismatches 11; Indels 1; Gaps 1;
QY 12 GNTTCGGCGGGCTGGTGGAGAAATGGAGCCGGTAGNTGCTTGGCGGAGTCCCGGNTCC 71
DB 85 GGTTCGGCGGGCTGGTGGAGAAATGGAGCCGGTAGGCTGCTGGCGGAGTCCCGGPTCC 144
QY 72 TCCGTAGACCCGGCGGA-NACCTTCGTTGAGTAACCTGGCGGAGTGGTGGAGCGGTG 130
DB 145 TCCGTAGACCCGGCGGAGCACCTTCGTTGAGTAACCTGGCGGAGTGGTGGAGCGGTG 204
QY 131 CTCACCTTCCTGCCCGCCAAAGCGTTCGTCGGGTGGCGTGCCTGCTTATGGAGG 190
DB 205 CTCACCTTCCTGCCCGCCAAAGCGTTCGTCGGGTGGCGTGCCTGCTTATGGAGG 264
QY 191 GAGTGTGGCGAGAGTATTGGGACCCATCGGAGCGTAACCTGGATCTCCGAGGCGCTG 250
DB 265 GAGTGTGGCGAGAGTATTGGGACCCATCGGAGCGTAACCTGGATCTCCGAGGCGCTG 324
QY 251 GCGGAGGCGCGCCACCTGGNGGGGCATT 278
DB 325 GCGGAGGCGCGCCACCTGGAGGGGCATT 352

RESULT 5

BG209864 871 bp mRNA linear EST 06-NOV-2000
LOCUS 601873563F1 NIH_MGC_54 Homo sapiens cDNA clone IMAGE:4097486 5',
DEFINITION mRNA sequence.

ACCESSION BG209864
VERSION BG209864.1 GI:11103450
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 871)
TITLE NTH-MGC http://mgc.nci.nih.gov/

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov

Tissue Procurement: ATCC
CDNA Library Preparation: CLONETECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM962 row: n column: 15
High quality sequence stop: 579.
Location/Qualifiers

FEATURES

source
1. .871
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4097486"
/tissue_type="from:chronic myelogenous leukemia"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_54"
/notes="Organ: bone marrow; Vector: pDNR-LIB (Clontech); Site1: SfII (ggccctcgcc); Site2: SfII (ggccattagggc); Double-stranded cDNA was prepared from cell line RNA. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCCATTATGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCGGCGCGCCGACATG-dT(30)BN-3', (where B=A, C, G and N=A, C, G, or T). Average insert size 1.75 kb (range 0.9-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."
BASE COUNT 182 a 211 c 196 t

Query Match 86.8%; Score 241.4; DB 10; Length 871;
Best Local Similarity 95.5%; Pred. No. 1.5e-46;
Matches 256; Conservative 0; Mismatches 11; Indels 1; Gaps 1;
Qy 12 GNTTCCGGCGGCTGGTGGAGATGGAGCGGTAGTGTCTGCGCGGAGTCCCGGNTCC 71
Db 7 GGTTCGGCGGGCTGGTGGAGATGGAGCGGTAGTGTCTGCGCGGAGTCCCGGNTCC 66
Qy 72 TCGGTAGACCCCGGGA-NACCTTCTGTGTGAGTAACCTGCGCGAGGTGTGGAGCGGTGTG 130
Db 67 TCGGTAGACCCCGGAGACCTTCTGTGTGAGTAACCTGCGCGAGGTGTGGAGCGGTGTG 126
Qy 131 CTACCTTCTCTCCCGCCAGGCGTCTCGCGGTGGCGGTGCTGCGCGGTATGGAGG 190
Db 127 CTACCTTCTCTCCCGCCAGGCGTCTCGCGGTGGCGGTGCTGCGCGGTATGGAGG 186
Qy 191 GAGTGTGTGCGCAGAGTATTGCGGACCCATCGGAGCGTAACTCGGATCTCCGAGGCGCTG 250
Db 187 GAGTGTGTGCGCAGAGTATTGCGGACCCATCGGAGCGTAACTCGGATCTCCGAGGCGCTG 246
Qy 251 GCGGAGCGCGCCACCTGGNGGGGCATT 278
Db 247 GCGGAGCGCGCCACCTGGAGGGGCATT 274

RESULT 6
AL519688
LOCUS 975 bp mRNA linear EST 22-MAY-2003
DEFINITION Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED Homo sapiens CDNA clone CS0DB004Y103 5-PRIME, mRNA sequence.
ACCESSION AL519688
VERSION AL519688.2 GI:31038034
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT On Feb 13, 2001 this sequence version replaced gi:12783181.
Contact: Genoscope
Genoscope Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 7864.f For more information about this cluster, see
http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DB004AE02QP1&cluster=7864.f. Contact :

Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DB004AE02QP1.
FEATURES
source 1. .975
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DB004Y103"
/tissue_type="NEUROBLASTOMA COT 10-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the PCWSPORT 6 vector. Library was normalized."
BASE COUNT 245 a 210 c 273 g 244 t 3 others
ORIGIN
Query Match 85.4%; Score 237.4; DB 9; Length 975;
Best Local Similarity 94.2%; Pred. No. 1.3e-45;
Matches 244; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
Qy 20 CGGCGTGGTGGAGATGGAGCGGTAGTGTCTGCGCGGAGTCCCGGNTCCCTCCGTAGA 79
Db 64 CGGCGTGGTGGAGATGGAGCGGTAGTGTCTGCGCGGAGTCCCGGNTCCCTCCGTAGA 123
Qy 80 CCGCGCGANACCTTCGTGTGAGTAACCTGCGCGGAGTGTGGAGCGGTGTGTCACCTTC 139
Db 124 CCGCGCGANACCTTCGTGTGAGTAACCTGCGCGGAGTGTGGAGCGGTGTGTCACCTTC 183
Qy 140 CTGCGCGCCCAAGCGGTGTCTGCGGTTGGCTGCGGTTGGCTGCGGTTGGCTGCGGTTGG 199
Db 184 CTGCGCGCCCAAGCGGTGTCTGCGGTTGGCTGCGGTTGGCTGCGGTTGGCTGCGGTTGG 243
Qy 200 CGCAGAGTATTGCGGACCCATCGGAGCGTAACTGGATCTCCGAGCGCTGGCGGAGGCC 259
Db 244 CGCAGAGTATTGCGGACCCATCGGAGCGTAACTGGATCTCCGAGCGCTGGCGGAGGCC 303
Qy 260 GGCACCTGNGGGGCATT 278
Db 304 GGCACCTGNGGGGCATT 322

RESULT 7
CB997897
LOCUS 770 bp mRNA linear EST 01-MAY-2003
DEFINITION AGENCOURT_13894636 NIH_MGC_148 Homo sapiens CDNA clone IMAGE:30348678 5', mRNA sequence.
ACCESSION CB997897
VERSION CB997897.1 GI:30292417
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: Dr. Stefan Hansson
CDNA Library Preparation: Michael J. Brownstein (NHGRI) with help and advice from Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM391 row: m column: 07
High quality sequence stop: 633.
Location/Qualifiers
1. .770

FEATURES
source

```
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30348678"
/tissue_type="pre-eclamptic placenta"
/lab_host="DH10B Tona"
/clone_lib="NIH_MGC_148"
/note="Organ: placenta; Vector: pBluescriptR; Site_1:
all-XhoI; Site_2: BamH; Library is oligo-dT primed and
directionally cloned using primer
5'-TTTTTTTTTTTTTTTNN-3', size-selected for average insert
size 2.3 kb and normalized to ROT 5. This is a primary
library enriched for full-length clones and constructed
using the Cap-trapper method (Carninci, in preparation).
Library constructed by M. Brownstein (NIMH/NHGRI,
National Institutes of Health). Note: this is a NIH_MGC
Library."
BASE COUNT      185 a   170 c   222 g   192 t      1 others
ORIGIN
Query Match      84.0%; Score 233.4; DB 14; Length 770;
Best Local Similarity 95.2%; Pred. No. 1.1e-44;
Matches 259; Conservative 0; Mismatches 11; Indels 2; Gaps 2;
QY 9 CTGNTTCGGCGGCTGGTGGAGGAGCGGTAGTCTTGGCGGAGTCCCG-GG 67
Db 25 CAGGGTTCCGGCGGCTGGTGGAGGAGCGGTAGTCTTGGCGGAGTCCCGGG 84
QY 68 NTCCTCCGTAGACCCCGGGA-NACCTTCGCTGTGAGTAACCTGCGGAGGTGGAGCG 126
Db 85 CTCCTCCGTAGACCCCGGAGACCTTCGTGTTGAGTAACCTGCGGAGGTGGAGCG 144
QY 127 TGTCTCACCTTCTGCGCCGCAAGGCGTTCGCGGGTGCCCTGCGCTGCGCTTATG 186
Db 145 TGTCTCACCTTCTGCGCCGCAAGGCGTTCGCGGGTGCCCTGCGCTGCGCTTATG 204
QY 187 GAGGAGTGTGCGCAGAGATTTCGGGACCCATCGGAGCGTAACTGGATCTCCGAGG 246
Db 205 GAGGAGTGTGCGCAGAGATTTCGGGACCCATCGGAGCGTAACTGGATCTCCGAGG 264
QY 247 CTTGGCGAGGCGCGCCACCTGGNGGGCATT 278
Db 265 CTTGGCGAGGCGCGCCACCTGGAGGGCATT 296
RESULT 8
CD110208      899 bp      mRNA      linear      EST 15-MAY-2003
LOCUS      AGENCOURT_13994860 NIH_MGC_148 Homo sapiens cDNA clone
DEFINITION      IMAGE:30331855 5', mRNA sequence.
ACCESSION      CD110208
VERSION      CD110208.1 GI:30754417
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 899)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Stefan Haussan
cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
and advice from Piero Carninci (RIKEN)
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM347 row: P column: 08
High quality sequence stop: 547.
```

```
location/Qualifiers
1. .899
'organism="Homo sapiens"
'mol_type="mRNA"
'db_xref="taxon:9606"
'clone="IMAGE:30331855"
'tissue_type="pre-eclamptic placenta"
'lab_host="DH10B Tona"
'clone_lib="NIH_MGC_148"
'note="Organ: placenta; Vector: pBluescriptR; Site_1:
all-XhoI; Site_2: BamH; Library is oligo-dT primed and
directionally cloned using primer
5'-TTTTTTTTTTTTTTTNN-3', size-selected for average insert
size 2.3 kb and normalized to ROT 5. This is a primary
library enriched for full-length clones and constructed
using the Cap-trapper method (Carninci, in preparation).
Library constructed by M. Brownstein (NIMH/NHGRI,
National Institutes of Health). Note: this is a NIH_MGC
Library."
BASE COUNT      218 a   199 c   261 g   221 t
ORIGIN
Query Match      84.0%; Score 233.4; DB 14; Length 899;
Best Local Similarity 95.2%; Pred. No. 1.1e-44;
Matches 259; Conservative 0; Mismatches 11; Indels 2; Gaps 2;
QY 9 CTGNTTCGGCGGCTGGTGGAGGAGCGGTAGTCTTGGCGGAGTCCCG-GG 67
Db 25 CAGGGTTCCGGCGGCTGGTGGAGGAGCGGTAGTCTTGGCGGAGTCCCGGG 84
QY 68 NTCCTCCGTAGACCCCGGGA-NACCTTCGCTGTGAGTAACCTGCGGAGGTGGAGCG 126
Db 85 CTCCTCCGTAGACCCCGGAGACCTTCGTGTTGAGTAACCTGCGGAGGTGGAGCG 144
QY 127 TGTCTCACCTTCTGCGCCGCAAGGCGTTCGCGGGTGCCCTGCGCTGCGCTTATG 186
Db 145 TGTCTCACCTTCTGCGCCGCAAGGCGTTCGCGGGTGCCCTGCGCTGCGCTTATG 204
QY 187 GAGGAGTGTGCGCAGAGATTTCGGGACCCATCGGAGCGTAACTGGATCTCCGAGG 246
Db 205 GAGGAGTGTGCGCAGAGATTTCGGGACCCATCGGAGCGTAACTGGATCTCCGAGG 264
QY 247 CTTGGCGAGGCGCGCCACCTGGNGGGCATT 278
Db 265 CTTGGCGAGGCGCGCCACCTGGAGGGCATT 296
RESULT 9
BX401591      1201 bp      mRNA      linear      EST 13-MAY-2003
LOCUS      BX401591 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
DEFINITION      Homo sapiens cDNA clone CS0DL005VE23 5-PRIME, mRNA sequence.
ACCESSION      BX401591
VERSION      BX401591.1 GI:30607193
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1201)
Li.W.B., Gruber,C., Jesse,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
7864.f For more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DL005AC120P1&cluster=7864.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
```

http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DL005AC12Qp1.

FEATURES

source

```
1. .1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DL005YE23"
/cell_line="RAMOS CELL LINE"
/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT
25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
```

BASE COUNT 287 a 235 c 328 g 304 t 47 others

Query Match 84.0%; Score 233.4; DB 13; Length 1201;
Best Local Similarity 93.7%; Pred. No. 1.1e-44;
Matches 251; Conservative 0; Mismatches 16; Indels 1; Gaps 1;

QY 12 GNTTCGGCGGCTGGTAGGAATGAGCCGCTAGTGTCTCGCGAGTCCCGGGTCC 71
DB 121 GGTTCGGCGGCTGGTAGGAATGAGCCGCTAGTGTCTCGCGAGTCCCGGGTCC 180

QY 72 TCCGTAGACCCCGGA-NACCTTCGTTGAGTAACCTGGCGAGTGGTGGAGGCTG 130
DB 181 TCCGTAGACCCCGGAGACCTTCGTTGAGTAACCTGGCGAGTGGTGGAGGCTG 240

QY 131 CTCACCTTCCTGCCCGCAAGCGTTGCTGGCGGTGGCTGCTGCTGCTGCTGAGG 190
DB 241 CTCACCTTCCTGCCCGCAAGCGTTGCTGGCGGTGGCTGCTGCTGCTGAGG 300

QY 191 GAGTGTGTCGCGAGAGTATTGGGACCATCGGAGCGTAACCTGGATCTCGCGAGCCTG 250
DB 301 GAGTGTGTCGCGAGAGTATTGGGACCATCGGAGCGTAACCTGGATCTCGCGAGCCTG 360

QY 251 GCGGAGCGCGCCACCTGGNGGGCATT 278
DB 361 GCGGAGCGCGCCACCTGGAGGGGCATT 388

RESULT 10

BI458126 679 bp mRNA linear EST 21-AUG-2001
LOCUS 603198548F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5278306 5',
DEFINITION mRNA sequence.
BI458126
ACCESSION BI458126.1 GI:15248782
VERSION 1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mhc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN).
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1702 row: o column: 11
High quality sequence stop: 679.
Location/Qualifiers

FEATURES

source

```
1. .679
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5278306"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/clone_lib="NIH_MGC_96"
/notes="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTWN-3',
size-selected for average insert size 2.3 kb and
normalized to 500 ng. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
```

BASE COUNT 166 a 152 c 198 g 163 t

Query Match 83.5%; Score 232; DB 12; Length 679;
Best Local Similarity 95.5%; Pred. No. 2.2e-44;
Matches 257; Conservative 0; Mismatches 10; Indels 2; Gaps 2;

QY 12 GNTTCGGCGGCTGGTAGGAATGAGCCGCTAGTGTCTCGCGAGTCCCGGGTCC 70
DB 10 GGTTCGGCGGCTGGTAGGAATGAGCCGCTAGTGTCTCGCGAGTCCCGGGTCC 69

QY 71 CTCGTAGACCCCGGA-NACCTTCGTTGAGTAACCTGGCGAGTGGTGGAGCGTGT 129
DB 70 CTCGTAGACCCCGGAGCACCTTCGTTGAGTAACCTGGCGAGTGGTGGAGCGTGT 129

QY 130 GCTCACCTTCCTGCCCGCAAGCGTTGCTGGCGGTGGCTGCTGCTGCTGAGG 189
DB 130 GCTCACCTTCCTGCCCGCAAGCGTTGCTGGCGGTGGCTGCTGCTGAGG 189

QY 190 GGAGTGTGTCGCGAGAGTATTGGGACCATCGGAGCGTAACCTGGATCTCGCGAGCCT 249
DB 190 GGAGTGTGTCGCGAGAGTATTGGGACCATCGGAGCGTAACCTGGATCTCGCGAGCCT 249

QY 250 GCGGAGCGCGCCACCTGGNGGGCATT 278
DB 250 GCGGAGCGCGCCACCTGGAGGGGCATT 278

RESULT 11

BI918213 736 bp mRNA linear EST 16-OCT-2001
LOCUS 603183167F1 NIH_MGC_121 Homo sapiens cDNA clone IMAGE:5246992 5',
DEFINITION mRNA sequence.
BI918213
ACCESSION BI918213.1 GI:16181971
VERSION 1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mhc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1623 row: f column: 17
High quality sequence stop: 736.
Location/Qualifiers

```

source
1. 736
/organism="Homo sapiens"
/mol_type="mrna"
/db_xref="taxon:9606"
/clone="IMAGE:5246992"
/lab_host="DHI0B"
/clone_lib="NIH_MGC_121"
/notes="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: EcoRV (destroyed); RNA source anonymous pool of 3
fetal brains, female age 20 weeks, female age 24 weeks,
and male age 26 weeks. Library is oligo-dr primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.7 kb, insert size range
0.7-3.5 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 017. Note:
this is a NIH_MGC Library."
BASE COUNT 173 a 172 c 215 g 176 t
ORIGIN
Query Match 83.5%; Score 232; DB 12; Length 736;
Best Local Similarity 95.5%; Pred. No. 2.3e-44;
Matches 257; Conservative 0; Mismatches 10; Indels 2; Gaps 2;
QY 12 GNTCCGGCGGGTGTGAGAAATGAGCGGTAGNTGCTTGGCGGAGTCCCG-GGNTC 70
Db 49 GGTTCGGCGGGTGTGAGAAATGAGCGGTAGNTGCTTGGCGGAGTCCCGGCGCTC 108
QY 71 CTCGGTAGACCCGGGA-NACCTTCGTTGAGTAACCTGGCGGAGGTGTGAGCGTGT 129
Db 109 CTCGGTAGACCCGGGAGACCTTCGTTGAGTAACCTGGCGGAGGTGTGAGCGTGT 168
QY 130 GCTCACCTTCTGCCCGCCAAAGCGTGTGCTGCGGGTGGCTGCTGCGCGCTTATGGAG 189
Db 169 GCTCACCTTCTGCCCGCCAAAGCGTGTGCTGCGGGTGGCTGCTGCGCGCTTATGGAG 228
QY 190 GGAGTGTGCGGAGAGTATGCGGAGCCCATCGGAGCGTAACCTGGATCTCCGAGCGCT 249
Db 229 GGAGTGTGCGGAGAGTATGCGGAGCCCATCGGAGCGTAACCTGGATCTCCGAGCGCT 288
QY 250 GCGGAGGCGCGCCACTGGNGGGGCATT 278
Db 289 GCGGAGGCGCGCCACTGGAGGGGCATT 317

RESULT 12
AUI43434 792 bp mrna linear EST 05-AUG-2002
LOCUS AUI43434 Y79AAL Homo sapiens cDNA clone Y79AAL001923 5', mRNA
DEFINITION
sequence.
ACCESSION AUI43434
VERSION AUI43434.1 GI:11004955
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 792)
AUTHORS Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
Yamanoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
Isogai,T.
TITLE HRI human cDNA project
JOURNAL Unpublished
COMMENT Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Iana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
Research Institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and

```

```

FEATURES
source
1. 792
/organism="Homo sapiens"
/mol_type="mrna"
/db_xref="taxon:9606"
/clone="Y79AAL001923"
/cell_type="retinoblastoma"
/cell_line="Y79"
/clone_lib="Y79AAL"
/notes="Vector: pME18SFL3"
BASE COUNT 183 a 180 c 235 g 191 t
ORIGIN
Query Match 83.5%; Score 232; DB 9; Length 792;
Best Local Similarity 95.5%; Pred. No. 2.3e-44;
Matches 257; Conservative 0; Mismatches 10; Indels 2; Gaps 2;
QY 12 GNTTCCGGCGGGTGTGAGAAATGAGCGGTAGNTGCTTGGCGGAGTCCCG-GGNTC 70
Db 65 GGTTCGGCGGGTGTGAGAAATGAGCGGTAGNTGCTTGGCGGAGTCCCGGCGCTC 124
QY 71 CTCGGTAGACCCGGGA-NACCTTCGTTGAGTAACCTGGCGGAGGTGTGAGCGTGT 129
Db 125 CTCGGTAGACCCGGGAGACCTTCGTTGAGTAACCTGGCGGAGGTGTGAGCGTGT 184
QY 130 GCTCACCTTCTGCCCGCCAAAGCGTGTGCTGCGGGTGGCTGCTGCGCGCTTATGGAG 189
Db 185 GCTCACCTTCTGCCCGCCAAAGCGTGTGCTGCGGGTGGCTGCTGCGCGCTTATGGAG 244
QY 190 GGAGTGTGTCGCGAGAGTATTGGCGGAGCCCATCGGAGCGTAACCTGGATCTCCGAGCGCT 249
Db 245 GGAGTGTGTCGCGAGAGTATTGGCGGAGCCCATCGGAGCGTAACCTGGATCTCCGAGCGCT 304
QY 250 GCGGAGGCGCGCCACTGGNGGGGCATT 278
Db 305 GCGGAGGCGCGCCACTGGAGGGGCATT 333

RESULT 13
BF691820 801 bp mrna linear EST 22-DEC-2000
LOCUS 602247789F1 NIH_MGC_62 Homo sapiens cDNA clone IMAGE:4332958 5',
DEFINITION mRNA sequence.
ACCESSION BF691820
VERSION BF691820.1 GI:11977228
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 801)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: ATCC/DCTD/DTMP
CDNA Library Preparation: CLONETECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LCH1200 row: a column: 23
High quality sequence stop: 665.
FEATURES
source
1. 801
/organism="Homo sapiens"
/mol_type="mrna"
/db_xref="taxon:9606"
/clone="IMAGE:4332958"
/issue_type="melanotic melanoma, high MDR"

```


THIS PAGE BLANK (usptd)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 28, 2003, 13:08:09 ; Search time 57 seconds
(without alignments)
2517.145 Million cell updates/sec

Title: US-10-042-417A-29

Perfect score: 515

Sequence: 1 cggtagtactgnttcggc.....cggccacctggngggcatt 278

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 1661050

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-O/cgn2_1/USPTO.spool/US10042417/runat_19082003_133538_9596/app_query.fasta_1.455
-DB-SPREMBL_23 -QFMT-fasten -SUFFIX-rspt -MINMATCH=0.1 -LOOPEXT=0
-UNITS-bits -START=1 -END=1 -MATRIX-BLOSUM62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=ptt -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10042417.ecgn.1.1.17-runat_19082003_133538_9596 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvrius:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	354	68.7	81	4 Q9UKCO	Q9ukc0 homo sapien

ID	Q9UKCO	PRELIMINARY:	PRT:	81 AA.
AC	Q9UKCO;			
DT	01-MAY-2000 (TrEMBLrel. 13, Created)			
DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)			
DT	01-JUN-2001 (TrEMBLrel. 17, Last annotation update)			
DE	F-box protein Fbx22 (Fragment).			
OS	FBX22.			
GN	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=200303060; PubMed=10531035;			
RA	Cenciarelli C., Chlaur D.S., Guardavaccaro D., Parks W., Vidal M.,			
RA	Pagano M.;			
RT	"Identification of a family of human F-box proteins."			
RL	Curr. Biol. 9:1177-1179(1999).			
DR	EMBL: AF174602; AAF04523.1;			
DR	InterPro: IPR001810; F-box.			
DR	Prfam; PF00646; F-box; 1.			
DR	SMART; SMO0256; FBOX; 1.			
FT	NON_TER 81			

ALIGNMENTS

RESULT 1

ID	Q9UKCO	PRELIMINARY:	PRT:	81 AA.
AC	Q9UKCO;			
DT	01-MAY-2000 (TrEMBLrel. 13, Created)			
DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)			
DT	01-JUN-2001 (TrEMBLrel. 17, Last annotation update)			
DE	F-box protein Fbx22 (Fragment).			
OS	FBX22.			
GN	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=200303060; PubMed=10531035;			
RA	Cenciarelli C., Chlaur D.S., Guardavaccaro D., Parks W., Vidal M.,			
RA	Pagano M.;			
RT	"Identification of a family of human F-box proteins."			
RL	Curr. Biol. 9:1177-1179(1999).			
DR	EMBL: AF174602; AAF04523.1;			
DR	InterPro: IPR001810; F-box.			
DR	Prfam; PF00646; F-box; 1.			
DR	SMART; SMO0256; FBOX; 1.			
FT	NON_TER 81			

Q9h824 homo sapien
Q8ne25 homo sapien
Q8w5k6 oryza sativ
Q8bv35 mus musculu
Q8vpr2 micrococci
Q9inx9 cercopithe
Q8udg5 agrobacteri
Q9w3g1 drosophila
Q91810 xenopus lae
Q95jcn sus scrofa
Q00600 homo sapien
Q8mqg8 caenorhabdi
Q02123 caenorhabdi
Q8mqg9 caenorhabdi
Q9af00 frankia sp.
Q9fcj3 streptomyce
Q9kzf9 streptomyce
Q95jd0 sus scrofa
Q95jdl sus scrofa
Q8kl17 mus musculu
Q9hgl0 emericella
Q9lgt8 oryza sativ
Q9gff2 streptomyce
Q8h3h1 oryza sativ
Q9p682 neurospora
Q07611 rattus norv
Q9kxg7 streptomyce
Q8qzy9 mus musculu
Q9fiq7 arabidopsis
Q9sx31 arabidopsis
Q94g66 oryza sativ
Q9k3u5 streptomyce
Q8qrv7 chimpanzee
Q94186 oryza sativ
Q04154 rattus norv
Q9gzhl caenorhabdi
P91019 caenorhabdi
Q8h8y1 oryza sativ
Q62105 mus musculu
Q8vcl9 mus musculu
Q9h5s6 homo sapien
Q8taj4 homo sapien
Q9jly2 mus musculu
Q8bpm0 mus musculu

SQ SEQUENCE 81 AA; 8768 MW; B8398FFC30C6CF4B CRC64;

Alignment Scores:

Pred. No.: 1.87e-29 Length: 81
 Score: 354.00 Matches: 71
 Percent Similarity: 86.59% Conservative: 0
 Best Local Similarity: 86.59% Mismatches: 10
 Query Match: 68.74% Indels: 1
 DB: 4 Gaps: 0

US-10-042-417a-29 (1-278) x Q9UKC0 (1-81)

QY 34 ATGGAGCCGCTAGTCTGCGCGAGTCCCGGCTCCCGTAGACCCCGGAGACCTT 93
 Db 1 MetGluProAlaGlyAlaCysGly-GlyProGlySerSerValAspProArgSerThrPh 20
 QY 94 CCGTTCAGTAACCTGCGGAGGTGGTGGAGCGTGTGCTCCTCCGCCCGCCCAAGGC 153
 Db 20 eValLeuSerAsnLeuAlaGluValValGluArgValLeuThrPheLeuProAlaLysAl 40
 QY 154 GTTGCTCGGCTGCGCTGCGCTTATGAGGAGGTGTGCGCAGAGTATTGGC 213
 Db 40 aLeuLeuArgValAlaCysValCysArgLeuTrpArgGluCysValArgValLeuAr 60
 QY 214 GACCATCGGAGCGTAACCTGATCTCCGCGAGCTGCGGAGCGCGCCACCTGGNGGG 273
 Db 60 gThrHisArgSerValThrTrpIleSerAlaGlyLeuAlaGluAlaGlyHisLeuGluGl 80
 QY 274 GCAT 277
 Db 80 yHis 81

RESULT 2

Q9H824
 ID Q9H824 PRELIMINARY; PRT; 276 AA.
 AC Q9H824;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical protein FLJ13986.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA ISOGAI T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
 RA Wagatsuma M., Hosoliri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
 RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
 RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
 RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,
 RA Ninomiya K., Iwayanagi T.,
 RA "NEDO human cDNA sequencing project."
 RT Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 RL EMBL; AK024048; BAB14798.1;
 DR InterPro; IPR001810; F-box.
 DR Pfam; PF00646; F-box; 1.
 DR SMART; SM00256; FBOX; 1.
 KW Hypothetical protein
 SQ SEQUENCE 276 AA; 30588 MW; 0E9B0CE3E208358B CRC64;

Alignment Scores:

Pred. No.: 6.78e-29 Length: 276
 Score: 349.00 Matches: 76
 Percent Similarity: 92.68% Conservative: 0
 Best Local Similarity: 92.68% Mismatches: 6
 Query Match: 67.77% Indels: 2
 DB: 4 Gaps: 0

US-10-042-417a-29 (1-278) x Q9H824 (1-276)

QY 34 ATGGAGCCGCTAGTCTGCGCGAGTCCCG-GGNTCTCCGTAGACCCCGGAGAN-ACC 91

Db 1 MetGluProValGlyCysGlyGluCysArgGlySerSerValAspProArgSerThr 20
 QY 92 TTCGTGTTGAGTAACCTGCGGAGGTGGTGGAGCGTGTGCTCACCCTTCCTGCCGCCCAAG 151
 Db 21 PheValLeuSerAsnLeuAlaGluValValGluArgValLeuThrPheLeuProAlaLys 40
 QY 152 GCGTTGCTGCGGCTGCGCTGCGCTTATGAGGAGGTGTGTCGCCGACAGATTG 211
 Db 41 AlaLeuLeuArgValAlaCysValCysArgLeuTrpArgGluCysValArgValLeu 60
 QY 212 CGGACCATCGGAGCGTAACCTGATCTCCGCGAGCTGCGGAGCGCGCCACCTCGNG 271
 Db 61 ArgThrHisArgSerValThrTrpIleSerAlaGlyLeuAlaGluAlaGlyHisLeuGlu 80
 QY 272 GGCAT 277
 Db 81 GlyHis 82

RESULT 3

Q8NEZ5
 ID Q8NEZ5 PRELIMINARY; PRT; 403 AA.
 AC Q8NEZ5;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE F-box protein FBX22p44 (F-box only protein 22).
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Tan P., Pan Z.-Q.;
 RT "FBX22p44: a novel human F-box protein predominantly expressed in the
 RT liver."
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Strausberg K.;
 RA Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
 RL EMBL; AY005144; AAF89095.1;
 DR EMBL; BC041691; AAH41691.1;
 DR Genew; HGNC:13593; FBXO22;
 DR InterPro; IPR001810; F-box.
 DR Pfam; PF00646; F-box; 1.
 SQ SEQUENCE 403 AA; 44508 MW; D96712BAA1149D8D CRC64;

Alignment Scores:

Pred. No.: 6.89e-29 Length: 403
 Score: 349.00 Matches: 76
 Percent Similarity: 92.68% Conservative: 0
 Best Local Similarity: 92.68% Mismatches: 6
 Query Match: 67.77% Indels: 2
 DB: 4 Gaps: 0

US-10-042-417a-29 (1-278) x Q8NEZ5 (1-403)

QY 34 ATGGAGCCGCTAGTCTGCGCGAGTCCCG-GGNTCTCCGTAGACCCCGGAGAN-ACC 91
 Db 1 MetGluProValGlyCysGlyGluCysArgGlySerSerValAspProArgSerThr 20
 QY 92 TTCGTGTTGAGTAACCTGCGGAGGTGGTGGAGCGTGTGCTCACCCTTCCTGCCGCCCAAG 151
 Db 21 PheValLeuSerAsnLeuAlaGluValValGluArgValLeuThrPheLeuProAlaLys 40
 QY 152 GCGTTGCTGCGGCTGCGCTGCGCTTATGAGGAGGTGTGTCGCCGACAGATTG 211
 Db 41 AlaLeuLeuArgValAlaCysValCysArgLeuTrpArgGluCysValArgValLeu 60
 QY 212 CGGACCATCGGAGCGTAACCTGGATCTCCGAGGCTGCGGAGCGCGCCACCTCGNG 271
 Db 81 GlyHis 82

```

Db 61 ArgThrHisArgSerValThrTrpIleSerAlaGlyLeuAlaGluAlaGlyHisLeuGlu 80
QY 272 GGCAT 277
DB 81 GlyHis 82
RESULT 4
Q8W5K6 PRELIMINARY; PRT; 1269 AA.
AC Q8W5K6;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 23, Last annotation update)
DE Hypothetical 138.4 kDa protein.
GN OSJNBA0079B05.10 OR OSJNAA0079B05.2.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA Wang R.A., Yu Y., Soderlund C., Chen M., Kim H.-R., Rambo T.,
RA Sasaki C., Henry D., Oates R., Simmons J.;
RT "Rice Genomic Sequence.";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Wang R.A., Yu Y., Yang T.J., Nah G., Soderlund C., Chen M., Kim H.-R.,
RA Rambo T., Sasaki C., Henry D., Oates R., Simmons J.;
RT "Rice Genomic Sequence.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC079179; AAL31655.1; -
DR EMBL; AC116601; AAM08709.1; -
DR Gramene; Q8W5K6; -
DR InterPro; IPR003104; FH2.
DR InterPro; IPR002965; P_rich_extensn.
DR Pfam; PF02181; FH2; 1.
DR PRINTS; PR01217; PRICHEXTENS.
DR SMART; SM00498; FH2; 1.
KW Hypothetical protein.
SQ SEQUENCE 1269 AA; 138432 MW; 262E546481B25CA6 CRC64;

Alignment Scores:
Pred. No.: 0.0505 Length: 1269
Score: 98.50 Matches: 30
Percent Similarity: 41.11% Conservative: 7
Best Local Similarity: 33.33% Mismatches: 32
Query Match: 20.18% Indels: 21
DB: 10 Gaps: 3

US-10-042-417A-29 (1-278) x Q8W5K6 (1-1269)
QY 278 AATGCCCNCCAGGTGGCCGCCCTCCCGAGGATCCAGGTTACGTCGCAT 219
DB 710 SerAlaProProProProProProProProProProAlaAsnArgSerAsnGlyProSerAla 729
QY 218 GGGTCCGCAATACTCTGGCGCACACACTCCCTCCATAAGCGGACGAGCCGCCGCA 159
DB 730 ProAlaProProProProProProProProProAlaAlaAsnLysArgAsnProProAla 749
QY 158 -----GCAACGCCTTGGCGGCGAGGAGTGAGCAGCAGCTCCACACCTCCGCGAGGT 105
DB 750 ProProProProProLeuMetThrGlyLysLysAlaProAlaProProProProProPro 768
QY 104 TACTCAACACGAAGTNTCCCGGGTCTACGAGGAGNCCCGGACTCCCGCGAGCA--- 48
DB 769 -----ProGlnAlaPro 772
QY 47 ---NCTACCGGCTCATTCCTCACCAGCCC 21
DB 773 LysProGlyThrValProProProPro 782

```

```

RESULT 5
Q8BV35 PRELIMINARY; PRT; 204 AA.
AC Q8BV35;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Hypothetical arginine-rich region containing protein (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Retina;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK080712; HAC37987.1; -
KW Hypothetical protein.
RT NON_TER 1
SQ SEQUENCE 204 AA; 20619 MW; 772B7C04FA90D4B2 CRC64;

Alignment Scores:
Pred. No.: 0.0867 Length: 204
Score: 96.00 Matches: 30
Percent Similarity: 37.50% Conservative: 6
Best Local Similarity: 31.25% Mismatches: 36
Query Match: 19.67% Indels: 24
DB: 11 Gaps: 3

US-10-042-417A-29 (1-278) x Q8BV35 (1-204)
QY 269 CCAGTGGCGGCGCTCCCGAGGATCCAGGTTACGTCGCATCCGCGCA 210
DB 101 ProSerGlyProProProProProProLeuArgSer----- 112
QY 209 ATACTCTGGCGCACACACTCCCTCCATAAGCGGACGAGCCGACGAGCCGACGCGCT 150
DB 113 -----GlyArgLeuProAlaAlaProGlyAspArgAlaGlyProCysAla---Pro 128
QY 149 TGGCGGCGAGGAGTGAGCAGCACACCTCCAGGATCCAGGTTACTCAACAGGAGG 90
DB 129 ArgAlaProSerArgProProHisAlaProGlySerAlaProProLeuAlaProProArg 148
QY 89 TNTCCGCGG-----GTCTACGAGGAGNCC 66
DB 149 AlaProThrProAlaValGlyGlyLeuGlyAlaThrLeuGluGlyValTrpGlyGlyHis 168
QY 65 CGGGACTCGCGCAGCAGCAGTCCAGGTCCTCCATTCCTCACCAGCCGCC 18
DB 169 AlaGluSerProArgLeuArgProGlyArgGlyProArgProArgThrAla 184

RESULT 6
Q8VPR2 PRELIMINARY; PRT; 216 AA.
AC Q8VPR2;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Putative collagen alpha 1 chain.
OS Micrococcus sp. 28.
OG Plasmid pSD10.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Micrococccineae; Micrococaceae; Micrococcus.
OX NCBI_TaxID=161213;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=28;
RA Zhong Z., Caspi R., Mincer T., Helinski D., Knauf V., Boardman K.,

```


QY	146	CGGCGAGCAAGGTGAGCACACGCTCCACCACCTCCGCCA-----GGTTACTCAACACGA	93
Db	134	AsnGlyGlyProArgTyr-TyrGluProProProProArgCysTyrPheGluAspArg	153
QY	92	AGGTNTCCGCGGGTCTACGGAGGA	69
Db	154	ArgValProAsnAlaTyrAspGly	161
RESULT	9		
Q9W3G1		PRELIMINARY;	PRT; 926 AA.
AC	Q9W3G1		
DT	01-MAY-2000	(TRENBLrel. 13, Created)	
DT	01-MAY-2000	(TRENBLrel. 13, Last sequence update)	
DE	01-OCT-2002	(TRENBLrel. 22, Last annotation update)	
DE	CG10555	protein.	
GN	CG10555.		
OS	Drosophila melanogaster (Fruit fly).		
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;		
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;		
OC	Ephydroidea; Drosophilidae; Drosophila.		
OX	NCBI_TaxID=7227;		
RP	[1]		
RN	SEQUENCE FROM N.A.		
RC	STRAIN=BERKELEY;		
RX	MEDLINE=20196006; PubMed=10731132;		
RA	Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,		
RA	Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,		
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,		
RA	Sutton G.G., Wortman J.R., Vandell M.D., Zhang Q., Chen L.X.,		
RA	Stratton R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,		
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,		
RA	Abriel J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,		
RA	Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,		
RA	Beeson K.Y., Bencs P.V., Berman B.P., Bhandari D., Bolshakov S.,		
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,		
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,		
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,		
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,		
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,		
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,		
RA	Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,		
RA	Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,		
RA	Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,		
RA	Hoslin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,		
RA	Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,		
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,		
RA	Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,		
RA	Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,		
RA	Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,		
RA	Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,		
RA	Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclev J.M.,		
RA	Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,		
RA	Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,		
RA	Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,		
RA	Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,		
RA	Swirskas R., Tector R., Turner R., Venter E., Wang A.H., Wang X.,		
RA	Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,		
RA	Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,		
RA	Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,		
RA	Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,		
RA	Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;		
RT	"The genome sequence of Drosophila melanogaster.;"		
RL	Science 287:2185-2193(2000).		
RL	EMBL; AE003443; AAF4366.1; -.		
DR	FlyBase; FBgn0030034; CG10555.		
DR	InterPro; IPR002965; P-rich.extensin.		
DR	PRINTS; PR01217; PRICHEXTENS.		
DR	SEQUENCE 926 AA. 93004 MW. 6FD5B9F77C36C006 CRC64;		

Alignment Scores:

0.134

Pred. No.:

926

Pred. No.:


```
Db 441 ---ProProProProAlaGly 446
RESULT 12
O00600 PRELIMINARY; PRT; 238 AA.
AC O00600;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Parotid 'o' protein (Fragment).
GN PRB4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96108975; PubMed=8554050;
RA Azen E.A., Oberger E., Fisher S., Prakobphol A., Niece R.L.;
RT "PRB1, PRB2, and PRB4 coded polymorphisms among human salivary
RT concanavalin-A binding, II-1, and Po proline-rich proteins.";
RL Am. J. Hum. Genet. 58:143-153(1996).
DR EMBL: S80916; AAB50687.2;
FT NON_TER 1
SQ SEQUENCE 238 AA; 24005 MW; 4F44E947FFF3A6C1 CRC64;

Alignment Scores:
Pred. No.: 0.183 Length: 238
Score: 93.00 Matches: 30
Percent Similarity: 40.66% Conservative: 7
Best Local Similarity: 32.97% Mismatches: 28
Query Match: 19.06% Indels: 26
DB: 4 Gaps: 6

US-10-042-417a-29 (1-278) x O00600 (1-238)
QY 272 CMCACAGT-----GCCGGCCTCCCGCAGGCTCGGAGATCCAGGTACGCTCC 222
Db 86 ProProGlyLysProGluGlyArgProProGlnGlyGlyAsnGlnSerGln----- 102
QY 221 GATGGTCCCATATCTCTCGGCACACACTCCCTCCATATAGCGGCACACGCCGCCACCC 162
Db 103 -----GlyProProHisProGlyLysProGluArgProPro 115
QY 161 GCACCAACGCTTGGCGGCAGGAAGTGAGCAGCAGCTCCACACCTCGCCAGGTAC 102
Db 116 Pro-----GlnGlyGlyAsnGlnSerHisArgProProProProGlyLys 131
QY 101 TCACACGAAGTNTCCGGGGTCTACGGAGGAGNCCGGGACTCGCGCAAGCANCTACC 42
Db 132 ProGluArgProProProGln-----GlyGlyAsnGlnSerGln----- 144
QY 41 GGCTCCATCTCTCAC-----CAGCCGCCGCGA 15
Db 145 GlyProProHisProGlyLysProGluGly 155

RESULT 13
Q8MQG8 PRELIMINARY; PRT; 518 AA.
AC Q8MQG8;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE PON-75 protein (corresponding sequence W03D2.1b).
GN W03D2.1 OR PON-75.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).

RX MEDLINE=99069613; PubMed=9851916;
RA Waterston R.;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Rohlffing T., Wohlmann P.;
RT "The sequence of 2n elegans cosmid W03D2.";
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF000298; AAM97960.1;
DR Wormpep; W03D2.1b; CE31729.
DR InterPro; IPR002955; P_Rich_extensn.
DR PRINTS; PR01217; PRICEXTENS.
SQ SEQUENCE 518 AA; 50473 MW; 4801F27D9663651E CRC64;

Alignment Scores:
Pred. No.: 0.214 Length: 518
Score: 92.50 Matches: 27
Percent Similarity: 35.63% Conservative: 4
Best Local Similarity: 31.03% Mismatches: 25
Query Match: 18.95% Indels: 31
DB: 5 Gaps: 3

US-10-042-417a-29 (1-278) x Q8MQG8 (1-518)
QY 275 GCCCNCACAGTGGCGGCGCTCGCCAGGCTCGGAGATCCAGGTACGCTCCGATGGG 216
Db 231 AlaProAlaGlySerProPro----- 239
QY 215 TCGCAATACTCTCGGCACACACTCCCTCCATAAGCGGCACACAGGCCACCCGCCAGCA 156
Db 240 -----ProProProProLysGly 245
QY 155 ACGCCITGCGGGCAGGAAGTGAGCAGCAGCTCCACACCTCGGCCAGGTACTCAACA 96
Db 246 SerProProLeuAlaGly-----SerGlySerProProProProAlaAlaGlySer 263
QY 95 CGAAGTNTCCGGGGTCTACGGAGGAGNCCGGGACTCGCGCAAGCANCTACCGGCTCC 36
Db 264 ProProProProArg-----ThrGlySerProProProProProProThrGlySer 279
QY 35 ATTCCTCACACAGCCGCCGCGA 15
Db 280 ProProProProProAlaGly 286

RESULT 14
O02123 PRELIMINARY; PRT; 524 AA.
AC O02123;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical 51.2 kDa protein.
GN W03D2.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
```

RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RA Rohlifing T., Wohldmann P.;
RT "The sequence of C. elegans cosmid W03D2.2";
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
[3]

RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RA Waterston R.;
RT "Direct Submission";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF000298; AAC48255.2; -;
DR Wormpep; W03D2.1a; CE28501.
DR InterPro; IPR002965; P_rich_extensn.
DR PRINTS; PR01217; PRICHEXTENS.
KW Hypothetical protein.
SQ SEQUENCE 524 AA; 51210 MW; 023D77EFFD172CC5 CRC64;

Alignment Scores:
Pred. No.: 0.214 Length: 524
Score: 92.50 Matches: 27
Percent Similarity: 35.63% Conservative: 4
Best Local Similarity: 31.03% Mismatches: 25
Query Match: 18.95% Indels: 31
DB: 5 Gaps: 3

US-10-042-417A-29 (1-278) x 002123 (1-524)

QY 275 GCCCCNCCAGGTGGCGGCTCCGCCAGGCTGGGGAGATCCAGGTTACGCTCCGATGGG 216
Db ||||||| ||| |||||||
QY 237 AlaProProAlaGlySerProProPro----- 245
QY 215 TCCGCAATACTCTGCGCACACACTCCCTCCATAGCGGCACACGCGGCCACCGCGCAGCA 156
Db |||||||
QY 246 -----ProProProProLysGly 251
QY 155 ACGCCTTGGGGGAGGAGGTGAGCACACGCTCCACACCTCCGCCAGGTTACTCAACA 96
Db ::||| - ||||| :: ::||| |||||||
QY 252 SerProProLeuAlaGly-----SerGlySerProProProProAlaAlaGlySer 269
QY 95 CGAAGTNTCCGGGGTCTACGAGGAGNCCGGGACTCGCGGCAAGCANTACCGGCTCC 36
Db |||||||
QY 270 ProProProProArg-----ThrGlySerProProProProProThrGlySer 285
QY 35 ATTCCTCACACGCCCGCGGA 15
Db ||| |||||||
QY 286 ProProProProAlaGly 292

RESULT 15

Q8MOG9
ID Q8MOG9 PRELIMINARY; PRT; 539 AA.
AC Q8MOG9;
DT 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE PQN-75 protein (corresponding sequence W03D2.1c).
GN W03D2.1 OR PQN-75.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA Waterston R.;
RT "Genome sequencing of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;

RA Rohlifing T., Wohldmann P.;
RT "The sequence of C. elegans cosmid W03D2.2";
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RA Waterston R.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF000298; AAM97961.1; -;
DR Wormpep; W03D2.1c; CE14506.
DR InterPro; IPR002965; P_rich_extensn.
DR PRINTS; PR01217; PRICHEXTENS.
SQ SEQUENCE 539 AA; 52555 MW; D3D3CFD8EF42CDB9 CRC64;

Alignment Scores:
Pred. No.: 0.214 Length: 539
Score: 92.50 Matches: 27
Percent Similarity: 35.63% Conservative: 4
Best Local Similarity: 31.03% Mismatches: 25
Query Match: 18.95% Indels: 31
DB: 5 Gaps: 3

US-10-042-417A-29 (1-278) x Q8MOG9 (1-539)

QY 275 GCCCCNCCAGGTGGCGGCTCCGCCAGGCTGGGGAGATCCAGGTTACGCTCCGATGGG 216
Db ||||||| ||| |||||||
QY 252 AlaProProAlaGlySerProProPro----- 260
QY 215 TCCGCAATACTCTGCGCACACACTCCCTCCATAGCGGCACACGCGGCCACCGCGCAGCA 156
Db |||||||
QY 261 -----ProProProProLysGly 266
QY 155 ACGCCTTGGGGGAGGAGGTGAGCACACGCTCCACACCTCCGCCAGGTTACTCAACA 96
Db ::||| - ||||| :: ::||| |||||||
QY 267 SerProProLeuAlaGly-----SerGlySerProProProProAlaAlaGlySer 284
QY 95 CGAAGTNTCCGGGGTCTACGAGGAGNCCGGGACTCGCGGCAAGCANTACCGGCTCC 36
Db |||||||
QY 285 ProProProProArg-----ThrGlySerProProProProProThrGlySer 300
QY 35 ATTCCTCACACGCCCGCGGA 15
Db ||| |||||||
QY 301 ProProProProAlaGly 307

Search completed: August 28, 2003, 13:27:24
Job time : 66 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 28, 2003, 11:46:20 ; Search time 15.5 seconds
(without alignments)
1686.893 Million cell updates/sec

Title: US-10-042-417a-29

Perfect score: 515

Sequence: 1 cgcgtactgnttcggc.....cggccacctggggggcatt 278

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 255726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlh
-Q/cgn2_1/USPTO_spool/US10042417/runat_19082003_133537_9564/app_query.fasta.1.455
-DB=SwissProt_41 -QFMT=fastcan -SUFFIX=rs -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10042417@cgn.1.1.12 @runat_19082003_133537_9564 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	349	67.8	403	1 FX22_HUMAN	Q8nez5 homo sapien
2	95.5	19.6	240	1 PRA_MYCTU	O53426 mycobacteri
3	93	19.1	276	1 PRPL_HUMAN	P10162 homo sapien
4	93	19.1	481	1 CBL2_MOUSE	O8k3m5 mus musculu
5	91	18.6	680	1 CA1A_HUMAN	Q03692 homo sapien
6	90.5	18.5	249	1 PRA_MYCLE	P41484 mycobacteri
7	89	18.2	1219	1 MYSS_YEAST	Q04439 saccharomyc
8	88.5	18.1	247	1 PRB4_HUMAN	P10163 homo sapien
9	88.5	18.1	3124	1 CA1C_CHICK	P13944 gallus gall
10	88	18.0	174	1 PRPP_HUMAN	P81489 homo sapien
11	88	18.0	424	1 S3BA_HUMAN	Q15427 homo sapien
12	87	17.8	296	1 PRP3_MOUSE	P05143 mus musculu
13	84.5	17.3	306	1 SDP2_MOUSE	O9gw55 homo sapien
14	84	17.2	1175	1 HCN4_RABIT	O9tv66 oryctolagus
15	84	17.2	1466	1 CA13_HUMAN	P02461 homo sapien
16	83	17.0	234	1 PRPM_HUMAN	P10161 homo sapien
17	82.5	16.9	467	1 CBPA_DICDI	P35085 dictyosteli
18	82.5	16.0	1584	1 BAIL_HUMAN	O14514 homo sapien

ALIGNMENTS

RESULT 1

ID	FX22_HUMAN	STANDARD	PRT	403 AA
AC	Q8NEZ5			
DT	15-SEP-2003 (Rel. 42, Created)			
DT	15-SEP-2003 (Rel. 42, Last sequence update)			
DE	F-box only protein 22 (F-box protein FBX22p44).			
GN	FBX22 OR FBX22			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Tan P., Pan Z.-O.;			
RT	"FBX22p44: a novel human F-box protein predominantly expressed in the			
RT	liver.";			
RL	Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Brain;			
RX	MEDLINE=22388257; PubMed=12477932;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,			
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,			
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,			
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,			
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,			
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,			
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,			
RA	Villaillon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,			
RA	Fahney J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,			
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,			
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,			
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,			
RA	Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,			
RA	Schmerch A., Schein J.E., Jones S.J.M., Marra M.A.;			
RT	"Generation and initial analysis of more than 15,000 full-length			
RT	human and mouse cDNA sequences.";			

P38486 canis famil
O43516 homo sapien
P14866 homo sapien
P05142 mus musculu
Q60847 mus musculu
Q10690 mycobacteri
Q99954 homo sapien
P71613 mycobacteri
P13889 rubella vir
P70560 rattus norv
Q91k4s arabidopsi
P97384 mus musculu
P02812 homo sapien
P04280 homo sapien
Q99946 homo sapien
P13238 drosophila
P30117 epstein-bar
P06425 human papil
P58840 homo sapien
P02810 homo sapien
Q35449 mus musculu
P13983 nicotiana t
P46582 caenorhabdi
Q13233 homo sapien
Q99715 homo sapien
P06729 homo sapien
Q13319 homo sapien

1 LEG3_CANFA
1 WAIP_HUMAN
1 ROL_HUMAN
1 PRP2_MOUSE
1 CA1C_MOUSE
1 YK82_MYCTU
1 PRL5_HUMAN
1 IF2_MYCTU
1 POLN_RUBVT
1 CA1C_RAT
1 OMP1_ARATH
1 ANXB_MOUSE
1 PRP2_HUMAN
1 PRP1_HUMAN
1 NG5_HUMAN
1 VTU2_DROME
1 YKR4_EBV
1 VE4_HPV08
1 ACRL_HUMAN
1 PRPC_HUMAN
1 NG5_MOUSE
1 EXTN_TOBAC
1 YLB8_CABEL
1 M3K1_HUMAN
1 CA1C_HUMAN
1 CD2_HUMAN
1 CD5S_HUMAN

1021
1495
3063
351
367

16.8
81.5
16.7
558
261
3119
16.6
81
16.6
721
16.5
80
16.4
134
16.3
79.5
16.3
2205
16.2
317
79
78.5
16.1
395
78.5
16.1
503
16.0
251
78
16.0
331
16.0
331
16.0
306
16.0
168
15.8
217
15.8
229
15.8
232
15.7
166
15.7
306
15.7
620
15.7
1021
15.7
1495
15.7
3063
15.7
351
15.6
367
15.6

```

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- FUNCTION: Probably recognizes and binds to some phosphorylated
CC proteins and promotes their ubiquitination and degradation.
CC -1- SUBUNIT: PART OF A SCF (SKP1-CULLIN-F-BOX) PROTEIN LIGASE COMPLEX
CC (BY SIMILARITY).
CC -1- TISSUE SPECIFICITY: Predominantly expressed in liver.
CC -1- SIMILARITY: Contains 1 F-box domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AY005144; AAF89095.1; -
CC EMBL: BC041691; AAH41691.1; -
CC Genew: HGNC:13593; FBXO22.
CC InterPro: IPR001810; F-box.
CC Pfam: PF00646; F-box; 1.
CC PROSITE: PS50181; FBXO; FALSE_NEG.
CC Ubl conjugation.
CC FT DOMAIN 21 67 F-BOX.
CC SEQUENCE 403 AA; 44508 MW; D96712BAA1149D8D CRC64;
CC -----
Alignment Scores:
Pred. No.: 6.77e-27 Length: 403
Score: 349.00 Matches: 76
Percent Similarity: 92.68% Conservative: 0
Best Local Similarity: 92.68% Mismatches: 6
Query Match: 67.77% Indels: 2
DB: 1 Gaps: 0

US-10-042-417a-29 (1-278) x FX22_HUMAN (1-403)
QY 34 ATGAGCCGGTAGTCTGCGCGAGTCCCG-GGNTCTCCGTAGACCGCGGAN-ACC 91
Db 1 MetGluProValGlyCysCysGlyGluCysArgGlySerValAspProArgSerThr 20
QY 92 TTCGTGTTGAGTAACCTGGCGGAGGTGGTGAGCGGTGCTACCTCTCTGCCGCCAAG 151
Db 21 PheValLeuSerAsnLeuAlaGluValGluValGluValGluValGluValGluVal 40
QY 152 GCCTTGTGCGGTGCGCTGCGTCCGCTTATGAGGAGGTGTGTGCGCAGAGTATTG 211
Db 41 AlaLeuArgValAlaCysValCysArgGluTyrArgGluCysValArgValLeu 60
QY 212 CGGACCCATCGGAGCGTAACCTGGATCTCGCGAGGCTGGCGGAGCGCGCCACCTGGNG 271
Db 61 ArgThrHisArgSerValThrTrpIleSerAlaGlyLeuAlaGluAlaGlyHisLeuGlu 80
QY 272 GGCGAT 277
Db 81 GlyHis 82

RESULT 2
PRA_MYCTU STANDARD; PRT; 240 AA.
AC O53426;
DT 30-MAY-2000 (Rel. 39, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Proline-rich antigen homolog.
GN PRA OR RV1078 OR MT1109 OR MTV017.31.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230;

```

```

RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Braham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holtroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Winn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: STRONG, TO M.LEPRAE PRA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AL021897; CAAL17194.1; -
CC EMBL: AE006991; AAK45365.1; -
CC PIR: D70894; D70894.
CC TIGR: MT1109; -
CC TubercuList; RV1078; -
KW Repeat; Complete proteome.
SQ SEQUENCE 240 AA; 25076 MW; C9B142E6CC1A2609 CRC64;
Alignment Scores:
Pred. No.: 0.0717 Length: 240
Score: 95.50 Matches: 31
Percent Similarity: 36.56% Conservative: 3
Best Local Similarity: 33.33% Mismatches: 28
Query Match: 19.57% Indels: 31
DB: 1 Gaps: 5

US-10-042-417a-29 (1-278) x PRA_MYCTU (1-240)
QY 272 CNCCAGGTGGCGG-----CCTCCGCCAGGCTCGGAGATCCAGTTA 228
Db 6 ProProGlyGlySerTyrProProProProProProGly----- 19
QY 227 CGCTCGGATGGTCCGCAATACTCTCGGCACACACTCTCTCCATAGCGGCACACGAGG 168
Db 20 -----ProSer---GlyGlyHisGluPro 26
QY 167 CCACCCGCGAGCAGCGCTTGGGGGCGAGGAGTGAGCAGCAGCTCCACCACCTCGGCCA 108
Db 27 ProProAlaAlaProProGlyGlySerGly-----TyrAlaProProProPro 43
QY 107 -----GGTTACTCAACACGAGGTNTCCGGGGTCTTACGAGGAGGCCGCGGAC 60
Db 44 SerSerGlySerGlyTyrProProProProProProProGlyGlyGlyAlaTyrProPro 63
QY 59 TCGCCGCAAGCACTACCGGCTCCATTCCTCACCAGCCC 21
Db 64 ProProProSerAlaGlyGlyTyrAlaProProProProPro 76

RESULT 3
PRA_MYCTU STANDARD; PRT; 276 AA.
ID PRPL_HUMAN

```

AC P10162; P02813;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Salivary proline-rich protein PO (Allele K) [Contains: Peptide P-D]
DE (Fragment)
GN PRB4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89121439; PubMed=3220251;
RA Lyons K.M., Stein J.H., Smithies O.;
RT "Many protein products from a few loci: assignment of human salivary
RT proline-rich proteins to specific loci";
RL Genetics 120:255-265(1988).
RN [2]
RP SEQUENCE OF 207-276.
RX MEDLINE=83186122; PubMed=6841349;
RA Saitoh E., Iseura S., Sanada K.;
RT "Complete amino acid sequence of a basic proline-rich peptide, P-D,
RT from human parotid saliva";
RL J. Biochem. 93:495-502(1983).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X07715; CAA30543.1; ALT_SEQ.
DR PIR: S03176; PIHUSD.
DR MIM: 168730; .
DR MIM: 180990; .
DR GO: GO:0005576; C:extracellular; NAS.
DR InterPro: IPR000637; AT hook.
DR PRINTS: PR00929; ATHOOK.
KW Repeat; Parotid gland; Multigene family.
FT NON_TER 1 276
FT CHAIN 207 276 PEPTIDE P-D.
SQ SEQUENCE 276 AA; 27816 MW; 9F494926C979441A CRC64;

Alignment Scores:
Pred. No.: 0.127 Length: 276
Score: 93.00 Matches: 30
Percent Similarity: 40.66% Conservative: 7
Best Local Similarity: 32.97% Mismatches: 28
Query Match: 19.06% Indels: 26
DB: 1 Gaps: 6

US-10-042-417a-29 (1-278) x PRPL_HUMAN (1-276)
Qy 272 CCNCCAGGT-----GGCCGGCTCCGCCAGGCTCGGAGATCCAGGTACGCTCC 222
Db 124 ProProGlyLysProGluGlyArgProProGlnGlyGlyAsnGlnSerGln----- 140
Qy 221 GATGGGTCGCGATACTCTCGGCACACACTCCCTCCATATAGCGGCACACGAGCCACCC 162
Db 141 -----GlyProProHisProGlyLysProGluArgProPro 153
Qy 161 GCACACAGCGCTTGGCGGAGGAGGTGAGCACACGCTCCACACCTCCGCCAGGTAC 102
Db 154 Pro-----GlnGlyGlyAsnGlnSerHisArgProProProProGlyLys 169
Qy 101 TCACACAGAGGTNTCCGGGGTCTACGGAGGAGNCCGGGACTCGCGCGAAGCANTACC 42
Db 170 ProGluArgProProProGln-----GlyGlyAsnGlnSerGln----- 182
Qy 41 GGCTCCATTCCTCAC-----CAGCCGCCCGGA 15

Db 183 GlyProProHisProGlyLysProGlyLysProGlyLysProGlyLys 193
RESULT 4
CBL2_MOUSE
ID CBL2_MOUSE STANDARD; PRT; 481 AA.
AC Q8K3M5;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cdk5 and abl enzyme substrate 2 (interactor with cdk3 2) (IK3-2).
GN CABLES2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Liver, and Spleen;
RX MEDLINE=21952370; PubMed=11955625;
RA Sato H., Nishimoto I., Matsuoka M.;
RT "IK3-2, a relative to ik3-1/cables, is associated with cdk3, cdk5, and
RT c-abl.";
RL Biochim. Biophys. Acta 1574:157-163(2002).
CC -!- FUNCTION: Unknown. Probably involved in G1-S cell cycle
CC transition.
CC -!- SUBUNIT: Binds to cdk3, cdk5 and c-abl. The C-terminal cyclin-box-
CC like region binds to cdk5.
CC -!- TISSUE SPECIFICITY: Widely expressed.
CC -!- SIMILARITY: BELONGS TO THE CYCLIN FAMILY.
CC -!- CAUTION: It is uncertain whether Met-1 or Met-6 is the initiator.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AY049712; MAL2171.1; .
DR MGD: MGI:2182335; Cables2.
DR GO: GO:0005515; F:protein binding activity; IPI.
DR Pfam: PF00134; cyclin; 1.
KW Cell cycle; Cell division; Cyclin.
FT DOMAIN 7 11 POLY-ALA.
FT DOMAIN 15 113 PRO-RICH.
SQ SEQUENCE 481 AA; 52710 MW; 00DAIC706578B1B2 CRC64;

Alignment Scores:
Pred. No.: 0.13 Length: 481
Score: 93.00 Matches: 29
Percent Similarity: 41.86% Conservative: 7
Best Local Similarity: 33.72% Mismatches: 36
Query Match: 19.06% Indels: 14
DB: 1 Gaps: 4

US-10-042-417a-29 (1-278) x CBL2_MOUSE (1-481)
Qy 269 CCAGTGGCGCGCTCCGCCAGGCTG-----CGAGATCCAGGTACGTCGATGGG 216
Db 26 ProAlaAlaArgAsnProAlaValProAlaValProArgArgGlyAspSerArgArgGln 45
Qy 215 TCCGCATACTCTCGGCACACACTCCCTCCATATAGCGGCACACGAGCCACCCGACCA 156
Db 46 AlaAlaLeuPhePheLeuAsnAsnIleSerLeuAspGly-----ArgProProSerLeu 63
Qy 155 ACGCCTTGGCGGCGAGGAAGGTGAGCACACCTCCACACCTCCGCCAGGTACTCAACA 96
Db 64 GlyPro-----GlyGlyGlyLysProAlaProProProProProProThrGluAla 80
Qy 95 CGAAGGTNTCCGGGGTCTACGGAGGAGNCCGGGACTCGCGCGAAGCANTACCAGGTCC 36
Db 111
Qy 111

Db 81 ArgGluAlaProAla-----ProProAlaProProGlyGly 93
 Qy 35 ATTCCTCACCAGCCGCC 18
 Db 94 LeuProGlyLeuProAla 99
 RESULT 5
 CALA_HUMAN
 ID CALA_HUMAN STANDARD; PRT; 680 AA.
 AC Q03692;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Collagen alpha 1(X) chain precursor.
 GN COL10A1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92109659; PubMed=1764025;
 RA Thomas J.T., Cresswell C.J., Rash B., Nicolai H., Jones T.,
 RA Solomon E., Grant M.E., Boot-Handford R.P.;
 RT "The human collagen X gene. Complete primary translated sequence and
 RT chromosomal localization.";
 RL Biochem. J. 280:617-623(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93012005; PubMed=1397333;
 RA Reichenberger E., Beier F., Luvalle P., Olsen B.R., von der Mark K.,
 RA Bertling W.M.;
 RT "Genomic organization and full-length cDNA sequence of human collagen
 RT X.";
 RL FEBS Lett. 311:305-310(1992).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX Beier F., Lammi M.B., von der Mark K.;
 RA Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Williams S.;
 RN Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE OF 52-680 FROM N.A.
 RX MEDLINE=92267014; PubMed=1587271;
 RA Apte S.S., Seldin M.F., Hayashi M., Olsen B.R.;
 RT "Cloning of the human and mouse type X collagen genes and mapping of
 RT the mouse type X collagen gene to chromosome 10.";
 RL Eur. J. Biochem. 206:217-224(1992).
 RN [6]
 RP SEQUENCE OF 561-666 FROM N.A.
 RX MEDLINE=91243838; PubMed=2037056;
 RA Apte S., Mattei M.-G., Olsen B.R.;
 RT "Cloning of human alpha 1(X) collagen DNA and localization of the
 RT COL10A1 gene to the q21-q22 region of human chromosome 6.";
 RL FEBS Lett. 282:393-396(1991).
 RN [7]
 RP SEQUENCE OF 547-655 FROM N.A.
 RX MEDLINE=92077285; PubMed=1743401;
 RA Reichenberger E., Aigner T., von der Mark K., Stoeck H., Bertling W.;
 RT "In situ hybridization studies on the expression of type X collagen
 RT in fetal human cartilage.";
 RL Dev. Biol. 148:562-572(1991).
 RN [8]
 RP REVIEW ON VARIANTS.
 RX MEDLINE=97255959; PubMed=9101290;
 RA Kuivaniemi H., Tromp G., Prockop D.J.;
 RT "Mutations in fibrillar collagens (types I, II, III, and XI), fibril-
 RT associated collagen (type IX), and network-forming collagen (type X)
 RT cause a spectrum of diseases of bone, cartilage, and blood vessels.";
 RL Hum. Mutat. 9:300-315(1997).
 RN [9]

RP VARIANTS SMCD ASP-598 AND PRO-614.
 RX MEDLINE=94136476; PubMed=8304336;
 RA Wallis G.A., Rash B., Sweetman W.A., Thomas J.T., Super M., Evans G.,
 RA Grant M.E., Boot-Handford R.P.;
 RT "Amino acid substitutions of conserved residues in the
 RT carboxyl-terminal domain of the alpha 1(X) chain of type X collagen
 RT occur in two unrelated families with metaphyseal chondrodysplasia
 RT type Schmid.";
 RL Am. J. Hum. Genet. 54:169-178(1994).
 RN [10]
 RP VARIANT SMCD ARG-591.
 RX MEDLINE=94272470; PubMed=8004099;
 RA McIntosh I., Abbott M.H., Warman M.L., Olsen B.R., Francomano C.A.;
 RT "Additional mutations of type X collagen confirm COL10A1 as the
 RT Schmid metaphyseal chondrodysplasia locus.";
 RL Hum. Mol. Genet. 3:303-307(1994).
 RN [11]
 RP VARIANT SMCD VAL-618.
 RX MEDLINE=95181449; PubMed=7876225;
 RA Chan D., Cole W.G., Rogers J.G., Bateman J.F.;
 RT "Type X collagen multimer assembly in vitro is prevented by a Gly618
 RT to Val mutation in the alpha 1(X) NCI domain resulting in Schmid
 RT metaphyseal chondrodysplasia.";
 RL J. Biol. Chem. 270:4558-4562(1995).
 RN [12]
 RP VARIANTS SMCD ARG-545; GLU-595; HIS-597; LYS-617; ARG-644 AND GLY-648.
 RX MEDLINE=95331767; PubMed=7607655;
 RA Bonaventure J., Chaminade F., Maroteaux P.;
 RT "Mutations in three subdomains of the carboxy-terminal region of
 RT collagen type X account for most of the Schmid metaphyseal
 RT dysplasias.";
 RL Hum. Genet. 96:58-64(1995).
 RN [13]
 RP VARIANT SMCD PRO-600.
 RX MEDLINE=96375754; PubMed=8782043;
 RA Wallis G.A., Rash B., Sykes B., Bonaventure J., Maroteaux P.,
 RA Zabel B., Wynne-Davies R., Grant M.E., Boot-Handford R.P.;
 RT "Mutations within the gene encoding the alpha 1 (X) chain of type X
 RT collagen (COL10A1) cause metaphyseal chondrodysplasia type Schmid but
 RT not several other forms of metaphyseal chondrodysplasia.";
 RL J. Med. Genet. 33:450-457(1996).
 RN [14]
 RP VARIANTS SMCD GLU-18 AND ARG-18.
 RX MEDLINE=97220591; PubMed=9067753;
 RA Ikegawa S., Nakamura K., Nagano A., Haga N., Nakamura Y.;
 RT "Mutations in the N-terminal globular domain of the type X collagen
 RT gene (COL10A1) in patients with Schmid metaphyseal
 RT chondrodysplasia.";
 RL Hum. Mutat. 9:131-135(1997).
 RN [15]
 RP VARIANTS SMD GLU-595.
 RX MEDLINE=99057503; PubMed=9837818;
 RA Ikegawa S., Nishimura G., Nagai T., Hasegawa T., Ohashi H.,
 RA Nakamura Y.;
 RT "Mutation of the type X collagen gene 'COL10A1' causes
 RT spondylometaphyseal dysplasia.";
 RL Am. J. Hum. Genet. 63:1659-1662(1998).
 RN [16]
 RP VARIANT SMCD CYS-597.
 RX MEDLINE=99069781; PubMed=9852679;
 RA Sawai H., Ida A., Nakata Y., Koyama K.;
 RT "Novel missense mutation resulting in the substitution of tyrosine by
 RT cysteine at codon 597 of the type X collagen gene associated with
 RT Schmid metaphyseal chondrodysplasia.";
 RL J. Hum. Genet. 43:259-261(1998).
 CC -!- FUNCTION: TYPE X COLLAGEN IS A PRODUCT OF HYPERTHROPHIC
 CC CHONDROCYTES AND HAS BEEN LOCALIZED TO PRESUMPTIVE
 CC MINERALIZATION ZONES OF HYALINE CARTILAGE.
 CC -!- SUBUNIT: Homotrimer.
 CC -!- PTM: Prolines are hydroxylated in some or all of the chains.
 CC -!- unit (G-X-Y) are hydroxylated in some or all of the chains.
 CC -!- DISEASE: DEFECTS IN COL10A1 ARE THE CAUSE OF SCHMID TYPE
 CC METAPHYSEAL CHONDRODYSPLASIA (SMCD), A DOMINANTLY INHERITED


```

CC EMBL; X65546; CAA46515.1; ..
DR EMBL; U15183; AAA63035.1; ..
DR EMBL; AL583925; CAC31911.1; ..
DR PIR; A41497; A41497.
DR Leproma; ML2395; -.
KW Antigen; Repeat; Complete proteome.
FT DOMAIN 46 85 4 X 10 AA TANDEM REPEATS OF [PV]-G-G-S-
FT Y-P-P-P-P-P.
FT REPEAT 46 55 1-1.
FT REPEAT 56 65 1-2.
FT REPEAT 66 75 1-3.
FT REPEAT 76 85 1-4 (APPROXIMATE).
FT DOMAIN 101 156 2 X 23 AA APPROXIMATE REPEATS.
FT REPEAT 101 123 2-1.
FT REPEAT 134 156 2-2.
FT CONFLICT 211 212 QL--> HV (IN REF. 1).
SO SEQUENCE 249 AA; 26295 MW; 28E56587E1570DA CRC64;

Alignment Scores:
Pred. No.: 0.224 Length: 249
Score: 90.50 Matches: 29
Percent Similarity: 36.96% Conservative: 5
Best Local Similarity: 31.52% Mismatches: 41
Query Match: 18.55% Indels: 17
DB: 1 Gaps: 4

US-10-042-417A-29 (1-278) x PRA_MYCLE (1-249)
QY 272 CCNCCAGTGGCGG-----CCTCCGCGAGCGCTCCGAGATCCAGG 231
Db 6 ProProSerGlySerAsnProThrProAlaProProProGly----- 20
QY 230 TTACGCTCCGATGGTCCGCAATACTCTGCGCACACACTCCCTCCATAACGCGCACAGC 171
Db 21 -----SerSerGlyGlyTyrGluProSerPheAlaProSerGluLeuGlySerAla 37
QY 170 AGGCCACCCGACGACGCTTGGCGGCGAGAGGTGACGACGCTCCACACCTCCG 111
Db 38 TyrProProThrAlaProProAlaGlyGly-----SerTyrProProProPro 55
QY 110 CCAGGTACTACACGAAAGTNTCCGGGTCTACGAGGANGCCGGGACTCGCGCAA 51
Db 56 ProGlyGlySerTyrProProProProProProProProProProProProProPro 74
QY 50 GCANCTACCGGCTCCATTCCTACCAGCGCGCGGA 15
Db 75 ProSerThrGlyAlaTyrAlaProProProProProProProProProProPro 86

RESULT 7
MY55_YEAST STANDARD; PRT; 1219 AA.
AC Q04439;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE MYosin-5 isoform.
GN MYO5 OR YMR109 OR YMR178.08.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RX PubMed=9169872;
RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,
RA Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
RA Jags K., Lye G., Moule S., Odell C., Pearson D., Rajandream M.A.,
RA Rice P., Skelton J., Walsh S., Whitehead S., Barrell B.G.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome
RT XIII."
RL Nature 387:90-93(1997).

```

```

RN SUBUNIT, AND MUTAGENESIS OF TRP-1123.
RX MEDLINE=21898311; PubMed=11901111;
RA Mochida J., Yamamoto T., Fujimura-Kamada K., Tanaka K.;
RT "The novel adaptor protein, Mtlp, and Vrp1p, a homolog of
RT Wiskott-Aldrich syndrome protein-interacting protein (WIP), may
RT antagonistically regulate type I myosins in Saccharomyces
RT cerevisiae".
RL Genetics 160:923-934(2002).
CC -1- SUBUNIT: Binds via its SH3 domain to Bbcl.
CC -1- SIMILARITY: Contains 1 myosin-like globular head domain.
CC -1- SIMILARITY: Contains 1 SH3 domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; 249702; CAA89745.1; ..
CC PIR; S54570; S54570.
CC HSP; P08799; IMND.
CC SGD; S0004715; MYO5.
CC GO; GO:0005857; C:actin cortical patch (sensu Saccharomyces); IDA.
CC GO; GO:0007047; P:cell wall organization and biogenesis; IMP.
CC GO; GO:0006897; P:endocytosis; IMP.
CC GO; GO:0006887; P:exocytosis; IMP.
CC GO; GO:0007121; P:polar budding; IMP.
CC GO; GO:0009651; P:salinity response; IGI.
CC InterPro; IPR000048; IQ_region.
CC InterPro; IPR001609; myosin_head.
CC InterPro; IPR001452; SH3.
CC Pfam; PF00612; IQ; 1.
CC Pfam; PF00663; myosin_head; 1.
CC Pfam; PF00018; SH3; 1.
CC PRINTS; PR00193; MYOSINHEAVY.
CC PRODOM; PD000355; myosin_head; 1.
CC PRODOM; PD000066; SH3; 1.
CC SMART; SM00242; MYSC; 1.
CC SMART; SM00326; SH3; 1.
CC PROSITE; PS00002; SH3; 1.
CC MYosin; Coiled coil; ATP-binding; SH3 domain.
FT DOMAIN 1 1085 1147 MYOSIN HEAD-LIKE.
FT DOMAIN 1190 1218 SH3.
FT NP_BIND 129 136 COILED COIL (POTENTIAL).
FT DOMAIN 1011 1016 ATP (POTENTIAL).
FT DOMAIN 1060 1063 POLY-PRO.
FT DOMAIN 1073 1081 POLY-ALA.
FT DOMAIN 1204 1218 POLY-PRO.
FT DOMAIN 1123 1123 ASP/GLU-RICH (ACIDIC).
FT MUTAGEN 1123 1123 W->S: ABOLISHES INTERACTION WITH Bbcl.
SQ SEQUENCE 1219 AA; 136898 MW; DFFB9EC16B61CD29 CRC64;

Alignment Scores:
Pred. No.: 0.335 Length: 1219
Score: 89.00 Matches: 23
Percent Similarity: 50.00% Conservative: 9
Best Local Similarity: 35.94% Mismatches: 18
Query Match: 18.24% Indels: 14
DB: 1 Gaps: 3

US-10-042-417A-29 (1-278) x MY55_YEAST (1-1219)
QY 257 CCTCCGCCAGCGCTCGCG-----AGATCCAGGTTCAGCTCCGATGGTCC 213
Db 1014 ProProGlyMetGlnAsnLysAlaAlaThrArgArgSerValProAsnProAlaSer 1033
QY 212 GCANFACTCTGGCGCACAC-----ACTCCCTCCATAGACGCGCACACGAGG 168
Db 1034 ThrLeuThrAlaSerGlnSerAsnAlaArgProSerProThrAlaAlaThrArgAla 1053

```

```

QY 167 CCACCCGAGCAGCCCTTGGGGGCA-----GGAAGTGAGCAGCAGCTCCA 120
Db 1054 ThrProAlaAlaThrProAlaAlaAlaMetGlySerGlyArgGlnAlaAsnIlePro 1073
QY 119 CCACCTCGGCCA 108
Db 1074 ProProProPro 1077

RESULT 8
PRB4_HUMAN
ID PRB4_HUMAN STANDARD; PRT: 247 AA.
AC P10163; P02813;
DT 21-JUL-1986 (rel. 01, Created)
DT 13-AUG-1987 (rel. 05, Last sequence update)
DT 28-FEB-2003 (rel. 41, Last annotation update)
DE Salivary proline-rich protein PO precursor (Allele S).
GN PRB4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85289325; PubMed=2993301;
RA Maeda N., Kim H.-S., Azen E.A., Smithies O.;
RT "Differential RNA splicing and post-translational cleavages in the
RT human salivary proline-rich protein gene system.";
RN J. Biol. Chem. 260:11123-11130(1985).
RL [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89121440; PubMed=2851479;
RA Lyons K.M., Stein J.H., Smithies O.;
RT "Length polymorphisms in human proline-rich protein genes generated
RT by intragenic unequal crossing over.";
RL Genetics 120:267-278(1988).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; K03207; AAA60188.1; -.
DR EMBL; X07882; CRA30729.1; -.
DR PIR; S03176; PIHUSD.
DR Genew; HGNC:9340; PRB4.
DR MIM; 168730; -.
DR MIM; 180990; -.
KW Repeat; Parotid gland; Multigene family; Glycoprotein; Signal.
FT SIGNAL 1 16
FT CHAIN 17 247 SALIVARY PROLINE-RICH PROTEIN PO.
FT CHAIN 17 39 PROTEIN N1.
FT CHAIN 40 177 GLYCOSYLATED PROTEIN A.
FT CHAIN 37 57 MISSING (IN REF. 2).
FT CONFLICT 218 218 D -> A (IN REF. 2).
SQ SEQUENCE 247 AA; 25108 MW; 6A1943E435161691 CRC64;

Alignment Scores:
Pred. No.: 0.352 Length: 247
Score: 88.50 Matches: 33
Percent Similarity: 37.96% Conservative: 8
Best Local Similarity: 30.56% Mismatches: 28
Query Match: 18.14% Indels: 39
DB: 1 Gaps: 7

US-10-042-417a-29 (1-278) x PRB4_HUMAN (1-247)
QY 272 CCNCCAGGT-----GGCCGGCCTCCGCGAGGCTGCGAGATCCAGG----- 231
Db 74 ProProGlyLysProGluGlyArgProGlnGlyGlyAsnGlnSerGlnGlyProPro 93

```

```

QY 230 -----TTACGGTCCGATGGGTCC 213
Db 94 ProHisProGlyLysProGluArgProProGlnGlyGlyAsnGlnSerGlnGly--- 112
QY 212 GCAATATCTCTGGGACACACTCCCTCCATAAGCGGCACACGAGCCCGCCGACGACG 153
Db 113 -----ThrProProProProGly-----LysProGluArgPro 123
QY 152 CCTTGGTGGCAGGAGGTGAGCACACGCTCCACACCTCCGCCGAGTTACTCAACACGA 93
Db 124 ProProGlnGlyGlyAsnGlnSerHisArgProProProProGlyLysProGluArg 143
QY 92 AGTNTCCGGGGTCTACGAGGAGNCCCGGGACTCGCGCAGCAGCAGTACCGGCTCCATT 33
Db 144 ProProProGln-----GlyGlyAsnGlnSerGln-----GlyProPro 156
QY 32 CCTCAC-----CAGCCCGCCGGA 15
Db 157 ProHisProGlyLysProGluGly 164

RESULT 9
CA1C_CHICK
ID CA1C_CHICK STANDARD; PRT: 3124 AA.
AC P13944; Q04509;
DT 01-JAN-1990 (rel. 13, Created)
DT 01-NOV-1997 (rel. 35, Last sequence update)
DT 15-SEP-2003 (rel. 42, Last annotation update)
DE Collagen alpha 1(XII) chain precursor (Fibrochimerin).
GN COL12A1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=White leghorn;
RX MEDLINE=92011862; PubMed=1918137;
RA Yamagata M., Yamada K.M., Yamada S.S., Shinomura T., Tanaka H.,
RA Nishida Y., Obara M., Kimata K.;
RT "The complete primary structure of type XII collagen shows a chimeric
RT molecule with reiterated fibronectin type III motifs, von Willebrand
RT factor A motifs, a domain homologous to a noncollagenous region of
RT type IX collagen, and short collagenous domains with an Arg-Gly-Asp
RT site.";
RL J. Cell Biol. 115:209-221(1991).
RN [2]
RP SEQUENCE OF 2456-3124 FROM N.A., AND SEQUENCE OF 2772-2794 AND
RP 2846-2873.
RX MEDLINE=90062079; PubMed=2584192;
RA Gordon M.K., Gerecke D.R., Dublet B., van der Rest M., Olsen B.R.;
RT "Type XII collagen. A large multidomain molecule with partial
RT homology to type IX collagen.";
RL J. Biol. Chem. 264:19772-19778(1989).
RN [3]
RP SEQUENCE OF 2960-3076 FROM N.A.
RX MEDLINE=87317590; PubMed=3476925;
RA Gordon M.K., Gerecke D.R., Olsen B.R.;
RT "Type XII collagen: distinct extracellular matrix component
RT discovered by cDNA cloning.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:6040-6044(1987).
RN [4]
RP SEQUENCE OF 1-1283 FROM N.A. (ISOFORM SHORT), AND ALTERNATIVE
RP SPLICING.
RC TISSUE=Embryo;
RX MEDLINE=93042014; PubMed=1420368;
RA Trueb J., Trueb B.;
RT "The two splice variants of collagen XII share a common 5' end.";
RL Biochim. Biophys. Acta 1171:97-98(1992).
RN [5]
RP ALTERNATIVE SPLICING.
RX MEDLINE=95370352; PubMed=7642694;
RA Koch M., Bohrmann B., Matthison M., Hagios C., Trueb B., Chiquet M.;

```


"Large and small splice variants of collagen XII: differential expression and ligand binding.";
 J. Cell Biol. 130:1005-1014(1995).
 CC -|- FUNCTION: TYPE XII COLLAGEN INTERACTS WITH TYPE I COLLAGEN-CONTAINING FIBRILS, THE COL1 DOMAIN COULD BE ASSOCIATED WITH THE SURFACE OF THE FIBRILS, AND THE COL2 AND NC3 DOMAINS MAY BE LOCALIZED IN THE PERIFIBRILLAR MATRIX.
 CC -|- SUBUNIT: TRIMER OF IDENTICAL CHAINS EACH CONTAINING 190 kDa OF NONTRIPLE-HELICAL SEQUENCES.
 CC -|- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing; Named Isoforms-2;
 CC Comment-The final tissue form of collagen XII may contain homotrimers of either isoform Long or isoform Short or any combination of isoform Long and isoform Short. Only isoform Long is a proteoglycan. Isoform Long has more restricted expression in embryonic tissue than isoform Short;
 CC Name-Long;
 CC IsoId-P13944-1; Sequence-Displayed;
 CC Name-Short;
 CC IsoId-P13944-2; Sequence-VSP_001148;
 CC LIGAMENTS, PERICHONDRUM, AND PERIOSTEUM, ALL DENSE CONNECTIVE TISSUES CONTAINING TYPE I COLLAGEN.
 CC -|- DOMAIN: THIS SEQUENCE DEFINES FIVE DISTINCT DOMAINS, TWO TRIPLE-HELICAL DOMAINS (COL1 AND COL2) AND THREE NONTRIPLE-HELICAL DOMAINS (NC1, NC2, AND NC3).
 CC -|- PTM: The triple-helical tail is stabilized by disulfide bonds at each end.
 CC -|- PTM: Prolines at the third position of the tripeptide repeating unit (G-X-Y) are hydroxylated in some or all of the chains.
 CC -|- PTM: O-glycosylated; glycosaminoglycan of chondroitin-sulfate type (By similarity).
 CC -|- SIMILARITY: BELONGS TO THE FIBRIL-ASSOCIATED COLLAGENS WITH INTERRUPTED HELICES (FACIT) FAMILY.
 CC -|- SIMILARITY: Contains 4 VWFA domains.
 CC -|- SIMILARITY: Contains 18 fibronectin type III domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; D00824; BAA00701.1; -;
 CC EMBL; X61024; CAA4358.1; -;
 CC EMBL; M17375; AAA48718.1; -;
 CC EMBL; J05137; AAA48635.1; -;
 CC EMBL; X67327; CAA47744.1; -;
 CC PIR; A40020; A40020.
 CC HSSP; P20701; 1LFA.
 CC InterPro; IPR000087; Collagen.
 CC InterPro; IPR003961; FN_III.
 CC InterPro; IPR003129; TSPN.
 CC InterPro; IPR002035; VWF_A.
 CC Pfam; PF01391; Collagen; 4.
 CC Pfam; PF00041; fn3; 17.
 CC Pfam; PF02210; TSPN; 1.
 CC Pfam; PF00092; vwa; 4.
 CC PRINTS; PR00453; VWFADOMAIN.
 CC SMART; SM00060; FN3; 16.
 CC SMART; SM00210; TSPN; 1.
 CC SMART; SM00327; VWA; 4.
 CC PROSITE; PS50234; VWFA; 4.
 CC Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
 CC Cell adhesion; Collagen; Signal; Glycoprotein; Alternative splicing.
 CC SIGNAL 1 24
 CC CHAIN 25 3124
 CC DOMAIN 25 114
 CC DOMAIN 139 311
 CC DOMAIN 332 425
 CC DOMAIN 439 615
 CC

FT DOMAIN 629 720
 FT DOMAIN 721 811
 FT DOMAIN 812 904
 FT DOMAIN 905 998
 FT DOMAIN 999 1085
 FT DOMAIN 1086 1178
 FT DOMAIN 1179 1371
 FT DOMAIN 1384 1473
 FT DOMAIN 1474 1565
 FT DOMAIN 1566 1654
 FT DOMAIN 1655 1755
 FT DOMAIN 1756 1846
 FT DOMAIN 1847 1936
 FT DOMAIN 1937 2027
 FT DOMAIN 2028 2118
 FT DOMAIN 2119 2206
 FT DOMAIN 2207 2295
 FT DOMAIN 2327 2500
 FT DOMAIN 2455 2750
 FT DOMAIN 2751 2902
 FT DOMAIN 2903 2945
 FT DOMAIN 2946 3048
 FT DOMAIN 3049 3124
 FT DOMAIN 3086 3096
 FT DOMAIN 3111 3123
 FT SITE 2438 2440
 FT SITE 2899 2901
 FT CARBOHYD 32 32
 FT CARBOHYD 797 797
 FT
 FT CARBOHYD 890 890
 FT CARBOHYD 981 981
 FT CARBOHYD 1006 1006
 FT CARBOHYD 1032 1032
 FT CARBOHYD 1044 1044
 FT CARBOHYD 1512 1512
 FT CARBOHYD 1767 1767
 FT CARBOHYD 2210 2210
 FT CARBOHYD 2273 2273
 FT CARBOHYD 2532 2532
 FT CARBOHYD 2683 2683
 FT VARSPLIC 25 1188
 FT CONFLICT 1258 1258
 FT CONFLICT 1264 1264
 FT CONFLICT 2759 2759
 FT CONFLICT 2803 2803
 FT CONFLICT 2977 2977
 FT CONFLICT 3075 3076
 SQ SEQUENCE 3124 AA; 340578 MW; 094285AFE7F346CF CRC64;
 Alignment Scores:
 Pred. No.: 0.389 Length: 3124
 Score: 88.50 Matches: 39
 Percent Similarity: 38.33% Conservative: 7
 Best Local Similarity: 32.50% Mismatches: 40
 Query Match: 18.14% Indels: 35
 DB: 1 Gaps: 7
 US-10-042-417A-29 (1-278) x CALIC_CHICK (1-3124)
 QY 269 CCAGGTGGCGCGCT-----CCGACAGCGCTGC---GGAGATCCAGTTAC-- 227
 Db 2965 proGlycylArgProGlyPheProGlyProGly-valGlnGlyProGlyGluAr 2984
 QY 227 ----- 227
 Db 2984 gGlyMetProGlyGluLysGlyGluArgGlyThrGlySerGlnGlyProArgGlyLeuPr 3004


```

Qy 226 -GCTCCGAGTGGTCCGCAATACTCTGCGCACACACTCCCTCCATA-----AGCGGCACACG 172
Db 3004 OGlyProGlyProGlnGlyGluSerArgThrGlyProGlySerThrGlySerAr 3024
Qy 171 CAGGCCACCGCA-----CAACGGCTTGGCGGAGAGGTGACACACGCTCCACC 118
Db 3024 GGLyProGlyProGlyProGlyArgProGlyAsnAlaGly-----lleargGlyProPr 3042
Qy 117 ACCTCCGCCAGGTTC-----TCAACACGAAGGTNTCCGGGTCTTACGAGGAGNCCCG 64
Db 3042 OGlyProGlyProGlyProGlyCysAspSerSerGlnCysAlaSerIleProTrpAsnGlyGlnG 3062
Qy 63 GGACTCCCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGTAC 6
Db 3062 yPheProGluProTrpValProGluSerGlyProTrpGlnProGluGlyGluProPhe 3081

RESULT 10
PRPP_HUMAN STANDARD; . PRT; 174 AA.
AC PRP1489;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Salivary proline-rich protein II-1 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RC TISSUE=Saliva;
RX MEDLINE=93385383; PubMed=8373986;
RA Kauffman D.L., Keller P.J., Bennick A., Blum M.;
RT "Alignment of amino acid and DNA sequence of human proline-rich
RT proteins.";
RL Crit. Rev. Oral Biol. Med. 4:287-292(1993).
DR GO: GO:0005576; C:extracellular; NAS.
DR InterPro: IPR000637; AT-hook.
DR PRINTS: PR00929; ATHOOK.
KW Repeat; Parotid gland; Multigene family.
FT NON_TER 174
SQ SEQUENCE 174 AA; 17802 MW; D645F106EB1BB5BE CRC64;

Alignment Scores:
Pred. No.: 0.389 Length: 174
Score: 88.00 Matches: 30
Percent Similarity: 37.78% Conservative: 4
Best Local Similarity: 33.33% Mismatches: 34
Query Match: 18.03% Indels: 22
DB: 1 Gaps: 5

US-10-042-417a-29 (1-278) x PRPP_HUMAN (1-174)
Qy 278 AATGCCCCCAGGTGGCGGCTCCGCCAGGCTCGGAGATCCAGTTACGCTCCGAT 219
Db 92 AsnGlnSerGlnGlyThrProProGlyLysProGluGlyArg----- 107
Qy 218 GGTCCGCAATACTCTCGGCACACACTCCCTCCATAGCGGCACACAGCCGCGCA 159
Db 108 -----ProGlnGlnGlyProGlyLysProGlu 118
Qy 158 GCACAGCTTGGCGGCGGAGAGTGAGACACAGCTCCACACCTCCGCGGAGTTACTCA 99
Db 119 ArgProProGlnGlyGlyAspGlnSerHisArgProProProGlyLysPro 138
Qy 98 ACACGAGGINTCCGGGGTCTACGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 39
Db 139 GluArgProProProGln-----GlyGly-----AspGlnSerGln-----Gly 151
Qy 38 TCCATTCTCTAC-----CAGCCCGCGCGGA 15
Db 152 ProProProHisProGlyLysProGluGly 161

```

```

RESULT 11
S3B4_HUMAN STANDARD; PRT; 424 AA.
AC Q15427;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 15-SEP-2003 (Rel. 42, Last annotation update)
DE Splicing factor 3B subunit 4 (Spliceosome associated protein 49) (SAP
DE 49) (SF3B50) (Pre-mRNA splicing factor SF3b 49 kDa subunit).
GN SF3B4 OR SAP49.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RC TISSUE=Skin;
RX MEDLINE=95047348; PubMed=7958871;
RA Champion-Arnaud F., Reed R.;
RT "The prespliceosome components SAP 49 and SAP 145 interact in a
RT complex implicated in tethering U2 snRNP to the branch site.";
RL Genes Dev. 8:1974-1983(1994).
RN [2]
RC TISSUE=Skin;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Lequellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McQuann P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RC CHARACTERIZATION OF THE SPLICEOSOME.
RX MEDLINE=20337962; PubMed=10882114;
RA Das R., Zhou Z., Reed R.;
RT "Functional association of U2 snRNP with the ATP-independent
RT spliceosomal complex E.";
RL Mol. Cell 5:779-787(2000).
CC -!- FUNCTION: SUBUNIT OF THE SPLICING FACTOR SF3B REQUIRED FOR 'A'
CC COMPLEX ASSEMBLY FORMED BY THE STABLE BINDING OF U2 snRNP TO THE
CC BRANCHPOINT SEQUENCE (BPS) IN PRE-MRNA. SEQUENCE INDEPENDENT
CC BINDING OF SF3A/SF3B COMPLEX UPSTREAM OF THE BRANCH SITE IS
CC ESSENTIAL, IT MAY ANCHOR U2 snRNP TO THE PRE-MRNA. MAY ALSO BE
CC INVOLVED IN THE ASSEMBLY OF THE 'E' COMPLEX. SF3B4 HAS BEEN
CC FOUND IN COMPLEX 'B' AND 'C' AS WELL, BELONGS ALSO TO THE
CC MINOR U12-DEPENDENT SPLICEOSOME, WHICH IS INVOLVED IN THE SPLICING
CC OF RARE CLASS OF NUCLEAR PRE-MRNA INTRON.
CC -!- SUBUNIT: COMPONENT OF SPLICING FACTOR SF3B WHICH IS COMPOSED OF
CC FOUR SUBUNITS: SF3B4/SAP49, SF3B3/SAP130, SF3B2/SAP145,
CC SF3B1/SAP155. SF3B ASSOCIATES WITH THE SPLICING FACTOR SF3A AND A
CC 12S RNA UNIT TO FORM THE U2 SMALL NUCLEAR RIBONUCLEOPROTEINS
CC COMPLEX (U2 snRNP). SF3B4 INTERACTS DIRECTLY WITH SF3B2.
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- SIMILARITY: STRONG, TO C.ELEGANS C08B11.5 AND, PARTIAL, TO
CC YEAST HSH49.
CC -!- SIMILARITY: Contains 2 RNA recognition motif (RRM) domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR	ENBL; L35013; AAA60300.1; -;	
DR	ENBL; BC004273; AAH04273.1; -;	
DR	ENBL; BC013886; AAH13886.1; -;	
DR	PIR; A54964; A54964.	
DR	HSSP; P11940; ICVJ.	
DR	Genew; HGNC:10771; SF3B4.	
DR	GK; Q15427; -;	
DR	MIM; 605593; -;	
DR	GO; GO:0005681; C:spliceosome complex; TAS.	
DR	GO; GO:0008248; F:pre-mRNA splicing factor activity; TAS.	
DR	GO; GO:0006371; P:mRNA splicing; TAS.	
DR	InterPro; IPR000504; RNA_rec_mot.	
DR	pfam; PF00076; rrm; 2.	
DR	SMART; SM00360; RRM; 2.	
DR	PROSITE; PS50102; RRM; 2.	
DR	PROSITE; PS00030; RRM_RNP_1; 1.	
KW	Spliceosome; mRNA processing; mRNA splicing; Nuclear protein;	
KW	RNA-binding; Repeat.	
FT	DOMAIN 13 91	RNA-BINDING (RRM) 1.
FT	DOMAIN 100 179	RNA-BINDING (RRM) 2.
FT	DOMAIN 215 218	POLY-PRO.
FT	DOMAIN 262 268	POLY-PRO.
SQ	SEQUENCE 424 AA; 444385 MW; 212472A25D3FF002 CRC64;	

Alignment Scores:		
Pred. No.:	0.403	Length:
Score:	88.00	Matches:
Percent Similarity:	41.67%	Conservative:
Best Local Similarity:	34.52%	Mismatches:
Query Match:	18.03%	Indels:
DB:	1	Gaps:
US-10-042-417A-29 (1-278) x S3B4_HUMAN (1-424)		
QY	272	CNCCAGGTGGCGCGCTCCGCCAGGCTCGGAGATCCAGGTTACGCTCCGATGGTCC
Db	297	ProProGlyGlyMetProHisProGlyMetSerGlnMetGlnLeuAlaHisGlyPro
QY	212	GCATACTCTGGGCACACT-----CCCTCCATAAGCGGCACAGCGAGCCACCC
Db	317	HisGlyLeu--GlyHisProHisAlaGlyProGlySerGlyGlyGlnProProPro
QY	161	GCAGCAACGCTTGGCGGGCAGGAGGTGAGCACGCTCCACCACCTCCGCCAGGTTAC
Db	336	ArgProProProGlyMet-----ProHisProGlyProProProMetGlyMet
QY	101	TCAACACGAAGTNTCCGCGGTCTACGGAGGAGNCCGGGACCTCGCGCAACNCTACC
Db	352	ProProArgGlyProPro-----PheGly-----SerProMetGlyHisPro
QY	41	GGCTCCATTCTCT
Db	366	GlyProMetPro

```

[1]
SEQUENCE FROM N.A.
RX MEDLINE=86059475; PubMed=2999141;
RA Ann D.K., Carlson D.M.;
RT "The structure and organization of a proline-rich protein gene of a
RL mouse multigene family.";
RT J. Biol. Chem. 260:15863-15872(1985).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M12100; AAA40005.1; -
CC MGD; MGI:97773; Prhl.
KW Repeat.
FT NON_TER
SQ SEQUENCE 296 AA; 29521 MW; 7F146824E8AF3269 CRC64;

Alignment Scores:
Pred. No.: 0.499 Length: 296
Score: 87.00 Matches: 36
Percent Similarity: 39.81% Conservative: 5
Best Local Similarity: 34.95% Mismatches: 47
Query Match: 17.83% Indels: 16
DB: 1 Gaps: 5

US-10-042-417A-29 (1-278) x PRP3_MOUSE (1-296)
QY 272 CCNCCAGGTGGCGGGCTCCGCCAGCGCTCGCGAGATCCA-----GGTTACGC 225
Db 195 ProProGlnGlyProProProGlnGlyGlyProHisProArgProThrGlnGlnGlyProPr 214
QY 224 TCCGATGGGTCCGCATACTCTGCGCACACACTCCCTCCATA-----AGCGGCACAG 172
Db 214 OProThrGlyProGlnProArgProThrGlnGlyProProProThrGlyGlyProGlnGln 234
QY 171 CAGCGCCACCGCCAGCAGCAACGCCCTTGGCGGCAGGA-----AGGTGAGCAGCAGCTCC 121
Db 234 nArGProProGlnGlyProProProProGlyGlyProGlnProArgProProGlnGlnGlyPr 254
QY 120 ACCACCTCCGCCAGGTCTACTCAACAGCA-----AGGTNTCCGGGCTCAGCGAGGANC 67
Db 254 OProProProThrGlyProGlnProArgProThrGlnGlyProHisProThrGlyGlyPr 274
QY 66 CCGGGACTCCGCCCAAGCANCT-----ACCGGCTCCATTCCTCACCAGCCGCCGG 16
Db 274 oGlnGlnThrProProLeuAlaGlyAsnProGlnGlyProProProGlyArgProGlnGln 294
QY 15 ANCCAG 9
Db 294 yProGln 296

RESULT 13
SDP2_HUMAN
ID SDP2_HUMAN STANDARD; PRT; 306 AA.
AC O9GZW5; Q9H9G1;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE SCAN domain-containing protein 2.
GN SCAND2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=20512041; PubMed=11056053;
RA Dupuy D., Aubert I., Duperat V.G., Petit J., Taine L., Stef M.,

```

RA Bloch B., Arveiler B.;
 RT "Mapping, characterization, and expression analysis of the SM-20
 RT human homologue, clorf12, and identification of a novel related gene,
 RT SCAND2";
 RL Genomics 69:348-354(2000).
 RN [2]
 RN SEQUENCE FROM N.A. (ISOFORM 2).
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,
 RA Matsunawa H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,
 RA Nakamura Y., Nagahara K., Masuho Y., Sasaki N.;
 RT NEDO human cDNA sequencing project.;
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RN SEQUENCE FROM N.A. (ISOFORM 2).
 RC TISSUE=Kidney, and Testis;
 RX MEDLINE=2238257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullighy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grinstead J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzyzinski M.I., Skalska U., Smalls D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=Q9GZW5-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q9GZW5-2; Sequence=VSP_007110, VSP_007111;
 CC -!- SIMILARITY: Contains 1 SCAN box domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF244812; AAC34567.1; -;
 DR EMBL; AF229246; AAC33966.1; -;
 DR EMBL; AK022844; BAB14268.1; -;
 DR EMBL; BC011547; AAH11547.1; -;
 DR EMBL; BC012929; AAH12929.1; -;
 DR EMBL; HGNC:10567; SCAND2.
 DR InterPro; IPR0003309; Treg_SCAN.
 DR Pfam; PF02023; SCAN; 1.
 DR SMART; SM00431; LER; 1.
 DR PROSITE; PS50804; SCAN_BOX; 1.
 KW Nuclear protein; Alternative splicing.
 FT DOMAIN 45 129 SCAN_BOX.
 FT DOMAIN 155 217 ALA-RICH.
 FT DOMAIN 191 276 ARG-RICH.
 FT VARSPPLIC 135 152 RARSGRTPPAQLRSPWP -> CLMWLEFLQRRAGVARR
 (in isoform 2)
 FT FTtd-VSP_007110.
 FT Missing (in isoform 2).
 FT FTtd-VSP_007111.

FT CONFLICT 94 P -> R (IN REF. 1).
 SQ SEQUENCE 306 AA; 34217 MW; D5E11136D30983EB CRC64;
 Alignment Scores:
 Pred. No.: 0.882 Length: 306
 Score: 84.50 Matches: 29
 Percent Similarity: 38.64% Conservative: 5
 Best Local Similarity: 32.95% Mismatches: 30
 Query Match: 17.32% Indels: 24
 DB: 1 Gaps: 4
 US-10-042-417a-29 (1-278) x SDP2_HUMAN (1-306)
 QY 276 TGCCCCN:CAGGTGG-----CCGGCCTCCGCCAGGCT 244
 ||| ||||| |||:|||||
 Db 173 CysAlaGlyArgTrpArgThrCysCysAlaAlaAlaAlaProSerAlaAlaArgSer 192
 QY 243 GCGGAGVCCAGGTGA---CGTCCGATGGTCCCACTACTCTCGGCACACACATCCCTC 187
 ||| :||| ||||| ||| |||||
 Db 193 AlaSer-AlaArgThrGlyArgSerThrSerCysAlaArgAlaAlaArgAlaProSe 212
 QY 186 CATAGC-----GGCACACGAGGCCA----- 165
 | :| | |||||
 Db 212 rAlaThrUluGlyAlaLeuThrArgThrProAlaProArgProLeuGlnArgArgAr 232
 QY 164 -CCGCAACACGCTTGGCGGAGGAGGTGAGCAGCAGCTCCACCACCTCCGCCAGG 106
 ||| ||||| ||||| ||||| |||||
 Db 232 gProGlyThrGlyProTrpArgProGlyArgGlnArgGlyAlaGlyThrAlaProProGl 252
 QY 105 TTACTCAACAGAGAGGTNTCCG 84
 | ||||| |||
 Db 252 yThrGlnProArgGlnArgPro 259
 RESULT 14
 HCN4_RABIT
 ID HCN4_RABIT STANDARD; PRT: 1175 AA.
 AC Q9TV66; Q9TVU35;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DE Potassium/sodium hyperpolarization-activated cyclic nucleotide-gated
 DE channel 4 (Hyperpolarization-activated cation channel 4) (HAC-4).
 GN HCN4 OR HAC4.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A., FUNCTION, AND TISSUE SPECIFICITY.
 RC TISSUE=Heart atrium;
 RX MEDLINE=99230313; PubMed=10212270;
 RA Ishii T.M., Takano M., Xie L.-H., Noma A., Ohmori H.;
 RT "Molecular characterization of the hyperpolarization-activated cation
 RT channel in rabbit heart sinoatrial node.";
 RL J. Biol. Chem. 274:12835-12839(1999).
 RN [2]
 RP SEQUENCE OF 261-381 FROM N.A.
 RC TISSUE=Heart;
 RX MEDLINE=99459217; PubMed=10400919;
 RA Shi W., Wymore R., Yu H., Wu J., Wymore R.T., Pan Z., Robinson R.B.,
 RA Dixon J.E., McKinnon D., Cohen I.S.;
 RT "Distribution and prevalence of hyperpolarization-activated cation
 RT channel (HCN) mRNA expression in cardiac tissues.";
 RL Circ. Res. 85:1-6(1999).
 CC -!- FUNCTION: Hyperpolarization-activated ion channel with very slow
 CC activation and inactivation exhibiting weak selectivity for
 CC potassium over sodium ions. May contribute to the native pacemaker
 CC currents in heart (If) and in neurons (Ih). Activated by cAMP. May
 CC mediate responses to sour stimuli.
 CC -!- SUBUNIT: The potassium channel is probably composed of a homo- or
 CC heterotetrameric complex of pore-forming subunits.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- TISSUE SPECIFICITY: Highly expressed in the heart sinoatrial node

RL Nucleic Acids Res. 16:2337-2337(1988).
 RN [8]
 RP REVISION TO 1184.
 RX MEDLINE-89098346; PubMed=3211760;
 RA Molyneux K., Dalglish R.;
 RT "Human type III collagen 'variant' is a cDNA cloning artefact.";
 RL Nucleic Acids Res. 16:11833-11833(1988).
 RN [9]
 RP SEQUENCE OF 1065-1466 FROM N.A.
 RX MEDLINE-85087944; PubMed=6096827;
 RA Loidl H.R., Brinker J.M., May M., Pihlajaniemi T., Morrow S.,
 RA Rosenbloom J., Myers J.C.;
 RT "Molecular cloning and carboxyl-propeptide analysis of human type III
 RT procollagen.";
 RL Nucleic Acids Res. 12:9383-9394(1984).
 RN [10]
 RP SEQUENCE OF 965-1200.
 RX MEDLINE-81208139; PubMed=7016180;
 RA Seyer J.M., Kang A.H.;
 RT "Covalent structure of collagen: amino acid sequence of alpha
 RT 1(III)-C89 from type III collagen of human liver.";
 RL Biochemistry 20:2621-2627(1981).
 RN [11]
 RP SEQUENCE OF 1176-1466 FROM N.A.
 RX MEDLINE-85157600; PubMed=2579949;
 RA Chu M.-L., Weil D., de Wet W.J., Bernard M.P., Sippola M., Ramirez F.;
 RT "Isolation of cDNA and genomic clones encoding human pro-alpha 1
 RT (III) collagen. Partial characterization of the 3' end region of the
 RT gene.";
 RL J. Biol. Chem. 260:4357-4363(1985).
 RN [12]
 RP SEQUENCE OF 1161-1200 FROM N.A.
 RX MEDLINE-86187804; PubMed=3754462;
 RA Miskulin M., Dalglish R., Kluge-Beckerman B., Rennard S.I.,
 RA Tolstoshev P., Brantly M., Crystal R.G.;
 RT "Human type III collagen gene expression is coordinately modulated
 RT with the type I collagen genes during fibroblast growth.";
 RL Biochemistry 25:1408-1413(1986).
 RN [13]
 RP SEQUENCE OF 1-170 FROM N.A.
 RX TISSUE-Placenta;
 RC MEDLINE-88303360; PubMed=3405773;
 RX Toman D., Ricca G., de Crombrughe B.;
 RT "Nucleotide sequence of a cDNA coding for the amino-terminal region
 RT of human prepro alpha 1(III) collagen.";
 RL Nucleic Acids Res. 16:7201-7201(1988).
 RN [14]
 RP SEQUENCE OF 1-176 FROM N.A.
 RX MEDLINE-89378752; PubMed=2777083;
 RA Benson-Chanda V., Su M.W., Weil D., Chu M.-L., Ramirez F.;
 RT "Cloning and analysis of the 5' portion of the human type-III
 RT procollagen gene (COL3A1).";
 RL Gene 78:255-265(1989).
 RN [15]
 RP REVIEW ON VARIANTS.
 RX MEDLINE-97255959; PubMed=9101290;
 RA Kuivaniemi H., Tromp G., Prockop D.J.;
 RT "Mutations in fibrillar collagens (types I, II, III, and XI), fibril-
 RT associated collagen (type IX), and network-forming collagen (type X)
 RT cause a spectrum of diseases of bone, cartilage, and blood vessels.";
 RL Hum. Mutat. 9:300-315(1997).
 RN [16]
 RP VARIANT AORTIC ANEURYSM ARG-303, AND VARIANT THR-668.
 RX MEDLINE-93293988; PubMed=8514866;
 RA Tromp G., Wu Y., Prockop D.J., Madhathari S.L., Kleinert C.,
 RA Earley J.J., Zhuang J., Noerregaard O., Darling R.C., Abbott W.M.,
 RA Cole C.W., Jaakkola P., Rynanen M., Pearce W.H., Yao J.S.T.,
 RA Majamaa K., Smellens S., Gatalica Z., Ferrell R.E., Jimenez S.A.,
 RA Jackson C.E., Michels V.V., Kaye M., Kuivaniemi H.;
 RT "Sequencing of cDNA from 50 unrelated patients reveals that mutations
 RT in the triple-helical domain of type III procollagen are an
 RT infrequent cause of aortic aneurysms.";
 RL J. Clin. Invest. 91:2539-2545(1993).

RN [17]
 RP VARIANT THR-698.
 RX MEDLINE-91045136; PubMed=2235526;
 RA Zafarullah K., Kleinert C., Tromp G., Kuivaniemi H., Kontusaari S.,
 RA Wu Y., Ganguly A., Prockop D.J.;
 RT "G to A polymorphism in exon 31 of the COL3A1 gene.";
 RL Nucleic Acids Res. 18:6180-6180(1990).
 RN [18]
 RP VARIANT AORTIC ANEURYSM ARG-786.
 RX MEDLINE-91056145; PubMed=2243125;
 RA Kontusaari S., Tromp G., Kuivaniemi H., Romanic A.M., Prockop D.J.;
 RT "A mutation in the gene for type III procollagen (COL3A1) in a family
 RT with aortic aneurysms.";
 RL J. Clin. Invest. 86:1465-1473(1990).
 RN [19]
 RP VARIANT EDS-IV ARG-828.
 RX MEDLINE-94016385; PubMed=8411057;
 RA Richards A.J., Narcisi P., Lloyd J.C., Fergusson C., Pope F.M.;
 RT "The substitution of glycine 661 by arginine in type III collagen
 RT produces mutant molecules with different thermal stabilities and
 RT causes Ehlers-Danlos syndrome type IV.";
 RL J. Med. Genet. 30:690-693(1993).
 RN [20]
 RP VARIANT EDS-IV SIR-957.
 RX MEDLINE-89109135; PubMed=2492273;
 RA Tromp G., Kuivaniemi H., Shikata H., Prockop D.J.;
 RT "A single base mutation that substitutes serine for glycine 790 of
 RT the alpha 1(III) chain of type III procollagen exposes an arginine
 RT and causes Ehlers-Danlos syndrome IV.";
 RL J. Biol. Chem. 264:1349-1352(1989).
 RN [21]
 RP VARIANT EDS-IV VAL-960.
 RX MEDLINE-95268429; PubMed=7749417;
 RA Tromp G., de Paete A., Nuytink L., Madhathari S.L., Kuivaniemi H.;
 RT "Substitution of valine for glycine 793 in type III procollagen in
 RT Ehlers-Danlos syndrome type IV.";
 RL Hum. Mutat. 5:179-181(1995).
 RN [22]
 RP VARIANT EDS-IV GIU-1014.
 RX MEDLINE-92316511; PubMed=1352273;
 RA Richards A.J., Ward P.N., Narcisi P., Nicholls A.C., Lloyd J.C.,
 RA Pope F.M.;
 RT "A single base mutation in the gene for type III collagen (COL3A1)
 RT converts glycine 847 to glutamic acid in a family with Ehlers-Danlos
 RT syndrome type IV. An unaffected family member is mosaic for the
 RT mutation.";
 RL Hum. Genet. 89:414-418(1992).
 RN [23]
 RP VARIANT EDS-IV ASP-1050.
 RX MEDLINE-90037070; PubMed=2808425;
 RA Tromp G., Kuivaniemi H., Stollie C.A., Pope F.M., Prockop D.J.;
 RT "Single base mutation in the type III procollagen gene that converts
 RT the codon for glycine 883 to aspartate in a mild variant of
 RT Ehlers-Danlos syndrome IV.";
 RL J. Biol. Chem. 264:19313-19317(1989).
 RN [24]
 RP VARIANT EDS-IV VAL-1077.
 RX MEDLINE-91374480; PubMed=1895316;
 RA Richards A.J., Lloyd J.C., Ward P.N., de Paete A., Narcisi P.,
 RA Pope F.M.;
 RT "Characterisation of a glycine to valine substitution at amino acid
 RT position 910 of the triple helical region of type III collagen in a
 RT patient with Ehlers-Danlos syndrome type IV.";
 RL J. Med. Genet. 28:458-463(1991).
 RN [25]
 RP VARIANT EDS-IV GLJ-1173.
 RX MEDLINE-93022543; PubMed=1357232;
 RA Johnson P.H., Richards A.J., Pope F.M., Hopkinson D.A.;
 RT Alignment Scores:
 Pred. No.: 1.05 Length: 1466
 Score: 84.00 Matches: 31
 Percent Similarity: 33.04% Conservative: 6

Best Local Similarity: 27.68% Mismatches: 31
Query Match: 17.21% Indels: 44
DB: 1 Gaps: 5

US-10-042-417A-29 (1-278) x CAL3_HUMAN (1-1466)

QY	269	CCAGGTGCGCGCTCCGCCAGGCTCGCGAGATCCAGGTTACGCTCCGATGGTCCGCA	210
Db	680	ProGlyGluArgGlyProProGlyLeu	688
QY	209	ATACTCTCGGCACACACTCCCTCCATAGCGGCACACGCGCCACCGCGCAGCAACGCCT	150
Db	689	-----AlaGlyAlaProGlyLeuArgGlyGlyAlaGlyProProGlyPro-----	703
QY	149	TGGCGGCGGAGGTGAGTACACACGCTCCACACCTCGCCAGGTTACTCA	99
Db	704	---GluGlyGlyLysGlyAlaAlaGlyProProGlyProProGlyAlaAlaGlyThrPro	722
QY	98	---ACACGAGGTNTCCCGGGTCTACGGAGGA	69
Db	723	GlyLeuGlnGlyMetProGlyGluArgGlyLeuGlySerProGlyProLysGlyAsp	742
QY	68	-----NCCCGGGACTCGCCCAAGCA	48
Db	743	LysGlyGluProGlyGlyProGlyAlaAspGlyValProGlyLysAspGlyProArgGly	762
QY	47	NCTACCGGCTCATT---CCTCACCAGCCGCCGGA	15
Db	763	ProThrGlyProLleGlyProProGlyProAlaGly	774

Search completed: August 28, 2003, 13:25:16
Job time : 24.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 28, 2003, 13:12:34 ; Search time 25 Seconds
(without alignments)

2138.791 Million cell updates/sec

Title: US-10-042-417A-29

Perfect score: 515

Sequence: 1 ccgtagtactgnttcggc.....cgccacctgnggggcatt 278

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 566616

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlh
-Q=/cgn2.1/USPTO.spool/US10042417/runat_19082003.133538.9629/app_query.fasta.1.455
-DB=pir.76 -QFWT=fastan -SUFFIX=rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10042417.ecgn.1.1.25_erunat_19082003.133538.9629 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	95.5	19.6	240	D70894	probable pra prote
C 2	95	19.5	167	AH2841	hypothetical prote
C 3	95	19.5	172	A97619	hypothetical prote
C 4	94	19.3	379	S31719	proline-rich prote
C 5	93	19.1	310	PIHUSD	salivary proline-r
C 6	92.5	19.0	539	T28770	hypothetical prote
C 7	91	18.6	680	CGHU1D	collagen alpha 1(X
C 8	90.5	18.5	249	A41497	36k antigen pra
C 9	89	18.2	1219	S54570	probable membrane
C 10	88.5	18.1	805	T49385	hypothetical prote
C 11	88.5	18.1	3124	A40020	collagen alpha 1(X
C 12	88	18.0	295	B48013	proline-rich prote
C 13	88	18.0	424	A54964	spliceosome-associ
C 14	87.5	17.9	347	S10571	mucin 1 precursor,

C 15	87	17.8	708	D96711	hypothetical prote
C 16	86	17.6	302	T32711	hypothetical prote
C 17	86	17.6	1357	T29265	hypothetical prote
C 18	85.5	17.5	278	B39066	proline-rich prote
C 19	85.5	17.5	301	E29149	proline-rich prote
C 20	85	17.4	403	S52796	prpL2 protein - hu
C 21	84	17.2	188	D29149	proline-rich prote
C 22	84	17.2	317	A28996	proline-rich prote
C 23	84	17.2	1466	1 CGHU7L	collagen alpha 1(I
C 24	83.5	17.1	574	T43556	wiskott-aldrich sy
C 25	83.5	17.1	574	T38819	wiskott-aldrich sy
C 26	83	17.0	240	B24264	proline-rich prote
C 27	83	17.0	2022	T48818	glucan 1,4-alpha-g
C 28	83	17.0	2715	T13049	evelid - fruit fly
C 29	82.5	16.0	319	F75420	hypothetical prote
C 30	82.5	16.9	883	A96662	hypothetical prote
C 31	82.5	16.0	1584	T00026	brain-specific ang
C 32	82	16.8	227	C29149	proline-rich prote
C 33	82	16.8	294	A49688	lactose-binding le
C 34	82	15.9	411	T47951	hypothetical prote
C 35	82	15.9	1007	T01437	hypothetical prote
C 36	81.5	16.7	129	T06174	pE40 protein - ba
C 37	81.5	16.7	558	A33616	heterogeneous ribo
C 38	81	16.6	177	S65780	glycine/proline-ri
C 39	81	16.6	240	A24284	proline-rich prote
C 40	81	16.6	497	T41015	proline rich prote
C 41	81	16.6	568	JC7317	cyclin-dependent k
C 42	81	16.6	2796	JC4743	fatty-acid synthas
C 43	80.5	16.5	300	SI9560	proline-rich prote
C 44	80.5	16.5	721	E70766	hypothetical prote
C 45	80	16.4	134	JC5572	proline-rich prote

ALIGNMENTS

RESULT 1

D70894

probable pra protein - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999

C:Accession: D70894

R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd,
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998

A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A>Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno
A:Reference number: A70500; MUID:98295987; PMID:9634230

A:Accession: D70894

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-240 <COL>

A:Cross-references: GB:AL021897; GB:AL123456; NID:g3256022; PIDN:CAAL17194.1; PID:el25

A:Experimental source: strain H37RV

C:Genetics:

A:Gene: pra

Alignment Scores:

Pred. No.:	0.0895	Length:	240
Score:	95.50	Matches:	31
Percent Similarity:	36.56%	Conservative:	3
Best Local Similarity:	33.33%	Mismatches:	28
Query Match:	19.57%	Indels:	31
DB:	2	Gaps:	5

US-10-042-417A-29 (1-278) x D70894 (1-240)

Oy 272 CCNCCAGGTGGCGG-----CCTCGCGAGCGCTCGGAGATCCAGGTGA 228

|||||

Db 6 ProProGlyGlySerTyProProProProProGly----- 19

|||||

Oy 227 CGTCCGATGGGTCCGCAATACTCTCGCGACACACTCCCTCCATAGCGGCACACGCGAGG 168

|||||

A:Reference number: S62891; MUID:96184506; PMID:8605987
A:Accession: S62891
A:Molecule type: protein
A:Residues: 241-252 <CHA>
A:Note: amino end of peptide designated basic proline-rich protein IB-5
A:Note: it is unclear from the peptide sequence whether this is a product of the PRB2
C:Genetics:
A:Gene: GDB:PRB4
A:Cross-references: GDB:I19514; OMIM:180990
A:Map position: 12p13.2-12p13.2
A:Introns: 22/1: 34/1
A:Note: the list of introns may be incomplete
C:Superfamily: proline-rich protein
C:Keywords: glycoprotein; saliva; tandem repeat
F:1-16/Domain: signal sequence #status predicted <SIG>
F:241-310/Product: proline-rich peptide P-D #status experimental <MAT>
F:66,87,171/Binding site: carboxydrate (Asn) (covalent) #status experimental
F:108,150,192,213,234/Binding site: carboxydrate (Asn) (covalent) #status predicted

Alignment Scores:
Pred. No.: 0.157 Length: 310
Score: 93.00 Matches: 30
Percent Similarity: 40.66% Conservative: 7
Best Local Similarity: 32.97% Mismatches: 28
Query Match: 19.06% Indels: 26
DB: 1 Gaps: 6

US-10-042-417A-29 (1-278) x PIHUSD (1-310)

QY	272	CNCCAGT-----	GGCGGCTCCGCCAGGCTCGGGAGATCCAGGTACGGTCC	222
Db	158	ProGlyLysProGluGlyArgProGlnGlyGlyAsnGlnSerGln	-----	174
QY	221	GATGGGTCCGAATACTCTCGGCACACACTCCCTCCATAGCGGCACACGCGGCACGCC	162	
Db	175	-----	GlyProProHisProGlyLysProGluArgProPro	187
QY	161	GCAGCAAGCCTTGGCGGGCAGGAAGTGAGCACAGCTCCACACACTCGCCAGGTTAC	102	
Db	188	Pro-----	GlnGlyGlyAsnGlnSerHisArgProProProGlyLys	203
QY	101	TCAACAGCAGGTNTCCGGCGGTCTACGGAGGAGNCCGGGACTCGCGGCAAGCANCTACC	42	
Db	204	ProGluArgProProGln-----	GlyGlyAsnGlnSerGln-----	216
QY	41	GGCTCCACTCTCTCAC-----	CAGCCGCGCGGA	15
Db	217	GlyProProHisProGlyLysProGluGly	227	

RESULT 6
T28770
hypothetical protein W03D2.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000
R:Rohlfing, T.; Wohldmann, P.
submitted to the EMBL Data Library, June 1997
A:Description: The sequence of C. elegans cosmid W03D2.
A:Reference number: Z20519
A:Accession: T28770
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-539 <ROH>
A:CROSS-references: EMBL:AF000298; PIDN:AAC48255.1; GSPDB:GN00022; CESP:W03D2.1
A:Experimental source: strain Bristol N2; clone W03D2
C:Genetics:
A:Gene: CESP:W03D2.1
A:Map position: 4
A:Introns: 40/3: 88/3; 115/3; 146/3; 173/3
C:Superfamily: collagen alpha 1(VIII) chain; complement C1q carboxyl-terminal homolog

Alignment Scores:
Pred. No.: 0.166 Length: 539

US-10-042-417A-29 (1-278) x S54570 (1-1219)

QY	257	CTCTCCGCCAGGCTGCGG-----AGATCCAGGTTACCTCCGATGGGTCC	213
Db	1014	ProProGlyMetGlnAsnLysAlaIaThrArgSerValProAsnProAlaSer	1033
		-----	-----
QY	212	GCAATACTCTGCGCACAC-----ACTCCCTCCATAACGCCACACACGAGG	168

167	CCACCCGCACACACGCTTGGCGGCA-----GGAAGGTGAGACACAGCTCCA	120
1054	ThrProAlaIaIaThrProAlaAlaAlaAlaMetGlySerGlyArgGlnAlaAsnIlePro	1073
119	CCACCTCCGCCA	108

DD
1074 P10P10P10P10 1077

T49385
 hypothetical protein B1D1.390 [imported] - *Neurospora crassa*
 C:Species: *Neurospora crassa*
 C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
 C:Accession: T49385
 R:Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; NY
 Submitted to the Protein Sequence Database, May 2000
 A:Reference number: Z25022
 A:Accession: T49385
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-805 <SCH>
 A:Cross-references: EMBL:AL355927; GSPDB:GN00116; NCSP:B1D1.390
 A:Experimental source: BAC clone B1D1; strain OR74A
 C:Genetics:
 A:Gene: NCSP:B1D1.390
 A:Map position: 6
 A:Introns: 54/3; 212/3

Alignment scores:					
Pred. No.:	0.409	Length:	805		
Score:	88.50	Matches:	33		
Percent Similarity:	34.29%	Conservative:	3		
Best Local Similarity:	31.43%	Mismatches:	32		
Query Match:	18.14%	Indels:	37		
DB:	2	Gaps:	5		
 US-10-042-417A-29 (1-278) x T49385 (1-805)					
QY	269	CGAGGTGG;CGGCCTCCGCCAGGCGTC-----	CGGAGATCC	234	
Db	363	ProGlyGlyArgProSerAlaLeuAlaValaIaGlyProSerAlaGlyAlaGlyGly	382		
QY	233	AGGTTAGC;TCCGAT-----GGTCGCGCAATACTCTGCCCA-----	198		
Db	383	AlaLeuSerIleaspSerLeuPheGlyGlnGlyAlaLeuAlaLeuAlaGlyAla	402		
QY	197	-----CACATCCTCCTAAAGCGGCACACGAGGCCACCCGAGCA	156		
Db	403	ThrArgLys;SerAlaThrProThrProSerGlnThrGlyThrProGlnProIleAla	422		
QY	155	ACGCCCTTG+CGGGCAGGAAGGTAGACACAGCTCCACCACCTCCGCCAGGTTACTCAACA	96		
Db	423	ThrPro-----IleProProPro-----	428		

QY	95	CGAAGGTTTCCTCGGGGTCATACGAGGAGGCCCGGAGCATCCGCGAAGGAGTACCGGCTCC	38
Db	429	---AlaValProProValAlaGlyAlaProThrProValProAlaAlaAlaGlySer	447
QY	35	ATTCTCTCACAGGCC	21
Db	448	LeuAlaAlaLeuPro	452

RESULT 11

A40020
collagen alpha 1(XII) chain precursor - chicken
N:Alternate names: fibrochimerin
C:Species: Gallus gallus (Chicken)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 19-Jan-2001
C:Accession: A40020; A34485; B34485; A28037; S22814; S22254; S28811
R:Yamagata, M.; Yamada, K.M.; Yamada, S.S.; Shinomura, T.; Tanaka, H.; Nishida, Y.; Obata, J. Cell Biol. 115, 209-221, 1991
A:Title: The complete primary structure of type XII collagen shows a chimeric molecule within region of type IX collagen, and short collagenous domains with an Arg-Gly-Asp site.
A:Reference number: A40020; MUID:92011862; PMID:1918137
A:Accession: A40020
A:Molecule type: mRNA
A:Residues: 1-3124 <VMA>
A:CROSS-references: GB:D00824; NID:922810; PIDN:BAA00701.1; PID:g222811
A:Note: In the authors' translation residues 1216-1219 are shown after residue 1235 and, R:Gordon, M.K.; Gerecke, D.R.; Dublet, B.; van der Rest, M.; Olsen, B.R.
J. Biol. Chem. 264, 19772-19778, 1989
A:Title: Type XII collagen. A large multidomain molecule with partial homology to type I
A:Reference number: A34485; MUID:90062079; PMID:2584192
A:Accession: A34485
A:Molecule type: mRNA
A:Residues: 2456-2758, 'A', 2760-2802, 'F', 2804-2976, 'F', 2978-3124 <GOR>
A:CROSS-references: EMBL:J05137; NID:g211284; PIDN:AAA48635.1; PID:g211285
A:Accession: B34485
A:Molecule type: protein
A:Residues: 2772-2792; 2846-2873 <GOR2>
R:Gordon, M.K.; Gerecke, D.R.; Olsen, B.R.
Proc. Natl. Acad. Sci. U.S.A. 84, 6040-6044, 1987
A:Title: Type XII collagen: distinct extracellular matrix component discovered by cDNA
A:Reference number: A28037; MUID:87317590; PMID:3476925
A:Accession: A28037
A:Molecule type: mRNA
A:Residues: 2960-2976, 'F', 2978-3074, 'AG' <GOR3>
A:CROSS-references: EMBL:M17375; NID:g211649; PIDN:AAA48718.1; PID:g211650
A:Note: this sequence has been revised in reference A34485
R:Koch, M.; Bernasconi, C.; Chiquet, M.
Eur. J. Biochem. 207, 847-856, 1992
A:Title: A major oligomeric fibroblast proteoglycan identified as a novel large form of
A:Reference number: S23814; MUID:92362621; PMID:1323460
A:Accession: S23814
A:Molecule type: protein
A:Residues: 'X', 1333, 'Q', 1335-1347; 1914-1928; 2504, 'X', 2506, 'X', 2508-2511, 'X', 2513-2517 <
R:Dublet, B.; van der Rest, M.
J. Biol. Chem. 262, 17724-17727, 1987
A:Title: Type XII collagen is expressed in embryonic chick tendons. Isolation of pepsin-
A:Reference number: S22254; MUID:88087065; PMID:3121603
A:Accession: S22254
A:Molecule type: protein
A:Residues: 2831-2832, 'T', 2834, 'R', 2836-2843; 3002-3014 <DUB>
R:Trueb, J.; Trueb, B.
Biochim. Biophys. Acta 1171, 97-98, 1992
A:Title: The two splice variants of collagen XII share a common 5' end.
A:Reference number: S28811; MUID:93042014; PMID:1420368
A:Accession: S28811
A:Molecule type: preliminary
A:Molecule type: mRNA
A:Residues: 1-24, 1189-1257, 'S', 1259-1263, 'E', 1265-1280 <TRU>
A:CROSS-references: EMBL:X67327
C:Genetics:
A:Introns: 2845/3; 2863/3; 2887/3; 2922/1; 2985/1; 3008/1; 3065/1
C:Superfamily: collagen alpha 1(XII) chain; fibronectin type III repeat homology; von Wille
C:Keywords: alternative splicing; cell binding; coiled coil; connective tissue; disulfid
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-3124/Product: collagen alpha 1(XII) chain #status predicted <WAT>
F:24, 1189-3124/Product: collagen alpha 1(XII) chain short splice form #status predicted
F:24-114/Domain: IIIA #status predicted <IIIA>
F:24-105/Domain: fibronectin type III repeat homology <FN3A>
F:137-301/Domain: von Willebrand factor type A repeat homology <VWA1>
F:332-425/Domain: IIIB #status predicted <IIIB>
F:332-414/Domain: fibronectin type III repeat homology <FN3B>
F:437-601/Domain: von Willebrand factor type A repeat homology <VWA2>

F:629-1178/Domain: IIIC #status predicted <IIIC>
F:630-711/Domain: fibronectin type III repeat homology <FN3C>
F:721-802/Domain: fibronectin type III repeat homology <FN3D>
F:812-895/Domain: fibronectin type III repeat homology <FN3E>
F:905-986/Domain: fibronectin type III repeat homology <FN3F>
F:995-1076/Domain: fibronectin type III repeat homology <FN3G>
F:1086-1169/Domain: fibronectin type III repeat homology <FN3H>
F:1197-1361/Domain: von Willebrand factor type A repeat homology <FN3I>
F:1384-2295/Domain: IIID #status predicted <IIID>
F:1384-1465/Domain: fibronectin type III repeat homology <FN3J>
F:1474-1557/Domain: fibronectin type III repeat homology <FN3K>
F:1566-1647/Domain: fibronectin type III repeat homology <FN3L>
F:1655-1738/Domain: fibronectin type III repeat homology <FN3M>
F:1756-1838/Domain: fibronectin type III repeat homology <FN3N>
F:1847-1928/Domain: fibronectin type III repeat homology <FN3O>
F:1937-2019/Domain: fibronectin type III repeat homology <FN3P>
F:2028-2110/Domain: fibronectin type III repeat homology <FN3Q>
F:2119-2199/Domain: fibronectin type III repeat homology <FN3R>
F:2207-2294/Domain: fibronectin type III repeat homology <FN3S>
F:2325-2490/Domain: von Willebrand factor type A repeat homology <FN3T>
F:2438-2440/Domain: cell adhesion #status predicted
F:2509-2750/Domain: IXP, homologous to NC4 domain of type IX collagen #status predicted
F:2751-2902/Domain: collagenous COL2 #status predicted <COL2>
F:2899-2901/Domain: cell attachment (R-G-D) motif
F:2903-2945/Domain: non-collagenous NC2 #status predicted <NC2>
F:2946-3048/Domain: collagenous COL1 #status predicted <COL1>
F:3049-3124/Domain: non-collagenous NC1 #status predicted <NC1>
F:32, 1006, 1032, 1044, 1512, 1767, 2210, 2273, 2532, 2683/Binding site: carbohydrate (Asn) (C
F:2780, 2789, 2836, 2842, 2860, 2866, 2869, 3004, 3007/Modified site: hydroxyproline (Pro) #S
Alignment Scores:
Pred. No.: 0.353 Length: 3124
Score: 88.50 Matches: 39
Percent Similarity: 38.33% Conservative: 7
Best Local Similarity: 32.50% Mismatches: 40
Query Match: 18.14% Indels: 35
DB: 1 Gaps: 7
US-10-042-417A-29 (1-278) x A40020 (1-3124)
QY 269 CCAGGTGGCGGCT-----CCGCCAGGCGCTGCG---GGAGATCCAGGTTAC-- 227
Db 2965 ProglyGlyArgProGlyPheProGlyProGly-VaIglnglyProGlyGluAr 2984
QY 227 ----- 227
Db 2984 gGlyMetProGlyGlyLysGlyGluArgGlyThrGlySerGlnGlyProArgGlyLeuPr 3004
QY 226 -GCTCCGATGGTCCGCAATACTCTGCGCACACTCCCTCCATPA-----AGCGGCACAGC 172
Db 3004 oGlyProGlyProGlyGlyGluSerArgThrGlyProGlySerThrGlySerAr 3024
QY 171 CAGGCCACCGCA-----CAACCGCTTGGCGGAGGAGGTGACACACAGCTCCACC 118
Db 3024 gGlyProGlyProGlyProGlyArgProGlyAsnAlaGly-----TleargGlyProPr 3042
QY 117 ACCTCCGCCAGGTTAC-----TCAACACCAAGGTNTCCGCGGTCTACGAGGANCCTG 64
Db 3042 oGlyProGlyTyrcysAspSerGlnCysAlaSerIleProTyrsnGlyGlnGl 3062
QY 63 GGACTCCGCCCAAGCANTACCGGCTCCATTCCTCCACAGCCGCGCGGAANCCAGTAC 6
Db 3062 yPheProGluProTyValProGluSerGlyProTyGlnProGluGlyGluProPhe 3081
RESULT 12
B48013
proline-rich proteoglycan 2 precursor, parotid - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 02-Jun-1995 #sequence_revision 02-Jun-1995 #text_change 20-Aug-1999
C:Accession: B48013
R:Castle, A.W.; Castle, J.D.
J. Biol. Chem. 268, 20490-20496, 1993
A:Title: Novel secretory proline-rich proteoglycans from rat parotid. Cloning and cha

A:Reference number: A48013; MUID:93388626; PMID:8376404
A:Accession: B48013
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-295 <CAS>
A:Cross-references: GB:L17318; NID:g310199; PIDN:AAA03074.1; PID:g310200
C:Superfamily: proline-rich protein
C:Keywords: extracellular protein; glycoprotein; tandem repeat

Alignment Scores:
Pred. No.: 0.515 Length: 295
Score: 88.00 Matches: 29
Percent Similarity: 36.84% Conservative: 6
Best Local Similarity: 30.53% Mismatches: 28
Query Match: 18.03% Indels: 32
DB: 2 Gaps: 5

US-10-042-417A-29 (1-278) x B48013 (1-295)

Qy 272 CCNCCAGTGGC-----CGGCTCCGCGCGCTGGGAGATCCAGGTACGGCTCC 222
|||||
Db 152 ProProGlyGlnArgProProGlnGly----- 163
Qy 221 GATGGTCCGATACTCTGGGACACACTCCCTCCATAGCGGCACACGAGGCCACGCC 162
|||||
Db 164 -----ProProGlnGlyGlyProGlnArgProPro 174
Qy 161 ---GCAGCAACGCCTTGGCGGAGGAAGTGGACACAGCTCCACCACTCGGCCAGGT 105
|||||
Db 175 GlnProGlySerPro-----GlnGlyProProProProGly 186
Qy 104 TACTCAACAGAGGTNTCCGCG-----GTCTACGGAGGAGCCGCGACTCCCG 54
|||||
Db 187 GlyProGlnGlnArgAlaProGlnGlyProProGlnGlyGlyProGlnArgProPro 206
Qy 53 CAAGCANCTACCGGCTCCATCTCTCACCAGCCCGCGGAANCCAG 9
|||||
Db 207 GlnProGlySerProGlnGlyProProProProGlyProGln 221

RESULT 13
A54964
spliceosome-associated protein SAP-49 - human
C:Species: Homo sapiens (man)
C:Date: 23-Mar-1995 #sequence_revision 05-Apr-1995 #text_change 24-Sep-1999
C:Accession: A54964
R:Champion-Arnaud, P.; Reed, R.
Genes Dev. 8, 1974-1983, 1994
A:Title: The prespliceosome components SAP 49 and SAP 145 interact in a complex implicat
A:Reference number: A54964; MUID:95047348; PMID:7958871
A:Accession: A54964
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-424 <CHA>
A:Cross-references: GB:L35013; NID:g556216; PIDN:AAA60300.1; PID:g556217
C:Superfamily: unassigned ribonucleoprotein repeat-containing proteins; ribonucleoprotei
F:14-81/Domain: ribonucleoprotein repeat homology <RRM4>
F:101-169/Domain: ribonucleoprotein repeat homology <RRM2>

Alignment Scores:
Pred. No.: 0.494 Length: 424
Score: 88.00 Matches: 29
Percent Similarity: 41.67% Conservative: 6
Best Local Similarity: 34.52% Mismatches: 35
Query Match: 18.03% Indels: 14
DB: 2 Gaps: 5

US-10-042-417A-29 (1-278) x A54964 (1-424)

Qy 272 CCNCCAGTGGCGGCTCCGCGCGCTGGGAGATCCAGGTACGGCTCCGATGGTCC 213
|||||
Db 297 ProProGlyGlyMetProHisProGlyMetSerGlnMetGlnLeuAlaHisHisGlyPro 316
Qy 212 GCAATACTCTGGCACACACT-----CCCTCCATAAGCGGCACACCGAGCCACCC 162

Db 317 HisGlyLeu---GlyHisProHisAlaGlyProProGlySerGlyGlnProProPro 335
|||||
Qy 161 GCACCAAGCGCTTGGCGGAGGAGGTGACACACGCTCCACACCTCCGCCAGGTAC 102
|||||
Db 336 ArgProProProGlyMet-----ProHisProGlyProProProMetGlyMet 351
|||||
Qy 101 TCACACAGAAAGGTNTCCGCGGCTACGAGGAGGAGCCGCGGACTCCGCGCAAGCANTACC 42
|||||
Db 352 ProProAlaGlyProPro-----PheGly-----SerProMetGlyHisPro 365
Qy 41 GGCTCCATCTCT 30
|||||
Db 366 GlyProMetPro 369

RESULT 14

S10571
mucin 1 precursor, secreted epithelial tumor antigen splice form - human
N:Contains: mucin 1 secreted breast-cancer-associated splice form
C:Species: Homo sapiens (man)
C:Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 01-Dec-2000
C:Accession: S10571; JN0100; I56024; S09706; S10217
R:Wreschner, D.H.; Hareuveni, M.; Tsarfaty, I.; Smorodinsky, N.; Horev, J.; Zaretsky, Eur. J. Biochem. 189, 463-473, 1990
A:Title: Human epithelial tumor antigen cDNA sequences. Differential splicing may gen
A:Reference number: S10571; MUID:90276413; PMID:2351132
A:Accession: S10571
A:Molecule type: mRNA
A:Residues: 1-347 <WRE>
A:Cross-references: EMBL:X52228; NID:g36434; PIDN:CAA36477.1; PID:g36435
R:Tsarfaty, I.; Hareuveni, M.; Horev, J.; Zaretsky, J.; Weiss, M.; Jeltsch, J.M.; Gar Gene 93, 313-318, 1990
A:Title: Isolation and characterization of an expressed hypervariable gene coding for
A:Reference number: JN0100; MUID:91033045; PMID:1688329
A:Accession: JN0100
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-135, 'Q', 137-142, 'E', 144-163, 204-208, 'A', 210-347 <TSA>
A:Cross-references: GB:M35093; NID:g182252; PIDN:AAB59612.1; PID:g182253
R:Xing, P.X.; Tjandra, J.J.; Reynolds, K.; McLaughlin, P.J.; Purcell, D.F.J.; Mckenzi J. Immunol. 142, 3503-3509, 1989
A:Title: Reactivity of anti-human milk fat globule antibodies with synthetic peptides
A:Reference number: I56024; MUID:89235154; PMID:2715633
A:Accession: I56024
A:Status: preliminary; translated from GB/EMBL/DBD
A:Molecule type: mRNA
A:Residues: 182-201 <RES>
A:Cross-references: GB:M26316; NID:g516622; PIDN:AAA36336.1; PID:g516623
R:Tendler, S.J.B.
Biochem. J. 267, 733-737, 1990
A:Title: Elements of secondary structure in a human epithelial mucin core peptide fra
A:Reference number: S09706; MUID:90253387; PMID:2339983
A:Accession: S09706
A:Molecule type: protein
A:Residues: 182-201 <TEN>
C:Genetics:
A:Gene: GDB:MUC1; PUM
A:Cross-references: GDB:I20705; OMIM:158340
A:Map position: lq21-lq23
A:Keywords: alternative splicing; tandem repeat
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-347/Product: mucin 1, secreted epithelial tumor antigen splice form #status pred
F:24-163,204-347/Product: mucin 1, secreted breast-cancer-associated splice form #sta

Alignment Scores:
Pred. No.: 0.569 Length: 347
Score: 87.50 Matches: 32
Percent Similarity: 35.29% Conservative: 4
Best Local Similarity: 31.37% Mismatches: 30
Query Match: 17.93% Indels: 36
DB: 2 Gaps: 5

US-10-042-417A-29 (1-278) x S10571 (1-347)

```
Qy 269 CCAGGTGCGCGCCTCCGCGAGCTCGCGGAGATCCAGGTTACGCTCCGATGGTGGCGCA 210
    ||| :||| ||| |||||
Db 122 ProasparnLysProAlaPro-----GlySerThr 131
Qy 209 ATACTTCGCGACAC---ACTCCCTCCATAAGCGGCACACGCGGCGCCGCGAGCAACG 153
    ||||| ||| ||||| ||||| :|||
Db 132 AlaProProAlaHisGlyValThrSerAlaProAspThrArgProProGlySerThr 151
Qy 152 CCTTGGCGGCGAGGAGTGAGCACAGCGCTCCACCA----- 117
    |||||
Db 152 -----AlaProProAlaHisGlyValThrSerAlaPro 162
Qy 116 -----CCTCGCGCAGGTTACTCAACACAGAGGNTCCGCGGGTCTACGGAGGAGGCC 66
    ||||| ||||| :||| |||
Db 163 AspThrArgProProGlySerThrAla-ProAlaAlaHisGlyValThrSerAlaPr 182
Qy 65 CGGACTCGCGCGCAAGCANCTACCGGCTCCATTCCTCACCAGCGCGCGGAANCCAGTAC 6
    ||||| ||||| ||||| :|||
Db 182 oAspThrArg-----ProAlaProGlySerThrAlaProProAlaHisGlyVa 198
Qy 5 TACG 2
    |||
Db 198 lThr 199
```

RESULT 15

```
D96711
hypothetical protein F24J5.8 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001. #text_change 31-Mar-2001
R:Accession: D96711
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: D96711
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-708 <STO>
A:Cross-references: GB:AE005173; NID:g5734709; PIDN:AD49974.1; GSPDB:GN00141
C:Genetics:
A:Gene: F24J5.8
A:Map position: 1
```

```
Alignment Scores:
Pred. No.: 0.592 Length: 708
Score: 87.00 Matches: 31
Percent Similarity: 40.22% Conservative: 6
Best Local Similarity: 33.70% Mismatches: 32
Query Match: 17.83% Indels: 23
DB: 2 Gaps: 4
```

US-10-042-417a-29 (1-278) x D96711 (1-708)

```
Qy 278 ATGCCCCNCAGGTGGCGCGCTCCGCGAGGCTCGGAGATCCAGGTTACGCTCCGAT 219
    |||:||||| ||||| |||
Db 12 AsnSerProValThrSerProProProLeu----- 23
Qy 218 GGTCTCCGCAATACTCTCGGCACACACTCCCTCCATAGCGGCGACACGCGCCACCGCA 159
    :||| ||| ||| :|||
Db 24 AsnAlaThrSerProAlaThrProProValThrSerProLeuProProSerAla 43
Qy 158 GCAACGCGCTTGGCGGCGAGGAGTGAGCACAGCTCCACACGCTCCGCGCGGTACTCA 99
    ||| ||| ||||| ||||| :|||
Db 44 ProProProAsnArg-----AlaProProProProProValThr 57
```

```
Qy 98 ACACGAAGGTNTCCGCGGGTCTACGGAGANCCGGGACTCGCGCGAAGCANCTACGGGC 39
    ||| ||| ||| |||
Db 58 ThrSerProProValAlaAsnGly-AlaPro-----ProPr 70
Qy 38 TCCATTCTCACCAGCCCGCGGAANCCAGTACT 5
    ||| ||| ||||| |||||
Db 70 oPro---LeuProLysProProGluSerSerSer 80
```

Search completed: August 28, 2003, 13:28:28

Job time : 34 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2003, 04:54:29 ; Search time 506 Seconds
(without alignments)

499.219 Million cell updates/sec

Title: US-10-042-417A-29

Perfect score: 278

Sequence: 1 ccgtagtactgnttcggc.....cggcaccctgnggggcatt 278

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 780815 seqs, 454324936 residues

Total number of hits satisfying chosen parameters: 1561630

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_NA_New.*

1: /cgn2_6/ptodata/2/pna/PCT_NEW_COMB.seq.*
2: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq.*
3: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq.*
4: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq.*
5: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq.*
6: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq.*
7: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	273	98.2	278	6	US-10-042-417A-29
2	232	83.5	1306	6	Sequence 29, Appl
3	213	76.6	1301	6	Sequence 2370, Ap
4	128.2	46.1	406	6	Sequence 402, Ap
5	113.6	40.9	2086	6	Sequence 2543, Ap
6	37.6	13.5	1471	6	Sequence 89, Appl
7	37.6	13.5	1566	6	Sequence 266, App
8	37.6	13.5	1566	6	Sequence 2233, Ap
9	37.6	13.5	1566	6	Sequence 2234, Ap
10	36.8	13.2	4813087	5	Sequence 265, App
11	36.6	13.2	1377	6	Sequence 75, Appl
12	35.8	12.9	11285	7	Sequence 266, App
13	35	12.6	955	6	Sequence 19636, A
14	35	12.6	1467	6	Sequence 2901, A
15	35	12.6	1467	6	Sequence 3069, Ap
16	33.8	12.2	1320	6	Sequence 3070, Ap
17	33.4	12.0	1847	6	Sequence 235, App
18	33.4	12.0	2096	6	Sequence 255, App
19	33.2	11.9	3732	1	Sequence 2899, Ap
20	33.2	11.7	14300	7	Sequence 57, Appl
21	32.6	11.7	277	6	Sequence 12093, A
22	32.2	11.6	840	6	Sequence 2808, Ap
23	32.2	11.6	2837	6	Sequence 270, App
24	32	11.5	2256646	6	Sequence 11, Appl
25	31.8	11.4	1600	6	Sequence 1, Appl1
26	31.6	11.4	350	6	Sequence 583, App
					Sequence 8930, Ap

c 27	31.6	11.4	1282	6	US-10-612-783-3347	Sequence 3347, Ap
c 28	31.2	11.2	972	6	US-10-294-433-494	Sequence 494, App
c 29	31.2	11.2	1233	6	US-10-115-831-62	Sequence 62, Appl
c 30	31.2	11.2	1233	6	US-10-294-433-62	Sequence 62, Appl
c 31	31.2	11.2	1281	6	US-10-105-837-849	Sequence 849, App
c 32	31.2	11.2	1281	6	US-10-286-897-1160	Sequence 1160, Ap
c 33	31.2	11.2	1281	6	US-10-258-898A-1160	Sequence 1160, Ap
c 34	31.2	11.2	1346	6	US-10-286-897-4732	Sequence 4732, Ap
c 35	31.2	11.2	1346	6	US-10-258-898A-4732	Sequence 4732, Ap
c 36	31.2	11.2	136593	7	US-60-487-610-19393	Sequence 19393, A
c 37	31	11.2	804	6	US-10-612-783-850	Sequence 850, App
c 38	31	11.2	16641	7	US-60-487-610-20029	Sequence 20029, A
c 39	30.8	11.1	690	6	US-10-273-573-1166	Sequence 1166, Ap
c 40	30.8	11.1	690	6	US-10-273-573-1368	Sequence 1368, Ap
c 41	30.8	11.1	96649	6	US-10-633-913-10	Sequence 10, Appl
c 42	30.6	11.0	1047	6	US-10-612-783-3141	Sequence 3141, Ap
c 43	30.6	11.0	1988	7	US-60-487-610-616	Sequence 616, App
c 44	30.6	11.0	2208	7	US-60-487-610-617	Sequence 617, App
c 45	30.6	11.0	6784	6	US-10-292-798-873	Sequence 873, App

ALIGNMENTS

RESULT 1

US-10-042-417A-29

; Sequence 29, Application US/10042417A

; GENERAL INFORMATION:

; APPLICANT: Pagano, M.

; TITLE OF INVENTION: METHODS TO IDENTIFY COMPOUNDS USEFUL FOR THE TREATMENT OF PROL

; FILE OF INVENTION: AND DIFFERENTIATIVE DISORDERS

; FILE REFERENCE: 5914-090-999

; CURRENT APPLICATION NUMBER: US/10/042,417A

; CURRENT FILING DATE: 2002-01-07

; PRIOR APPLICATION NUMBER: 60/260,179

; PRIOR FILING DATE: 2001-01-05

; NUMBER OF SEQ ID NOS: 92

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 29

; LENGTH: 278

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: modified base

; LOCATION: all n positions

; OTHER INFORMATION: n=a, c, g or t

US-10-042-417A-29

Query Match 98.2%; Score 273; DB 6; Length 278;
Best Local Similarity 100.0%; Pred. No. 5.8e-63;
Matches 278; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	CCGTAGTACTGNTTCCCGGGGGTGGTGAGGAATGGAGCCGGTAGTGTGCGGCGAG	60
DB	1	CCGTAGTACTGNTTCCCGGGGGTGGTGAGGAATGGAGCCGGTAGTGTGCGGCGAG	60
QY	61	TCCCGGGTCTCCCTAGACCGCGGANACCTTCCTGTGTGAGTAACCTGGCGAGGTGT	120
DB	61	TCCCGGGTCTCCCTAGACCGCGGANACCTTCCTGTGTGAGTAACCTGGCGAGGTGT	120
QY	121	GGAGCGTGTGCTACCTTCCTCCCGCCCAAGCGTTGCTGCGGGTGGCGTGGCGG	180
DB	121	GGAGCGTGTGCTACCTTCCTCCCGCCCAAGCGTTGCTGCGGGTGGCGTGGCGG	180
QY	181	CTTATGAGGAGTGTGCGCAGAGTATTCGGGACCCATCGGAGCGTAACCTGATCTC	240
DB	181	CTTATGAGGAGTGTGCGCAGAGTATTCGGGACCCATCGGAGCGTAACCTGATCTC	240
QY	241	CGCAGGCCGCGGAGCGCGCCACCTGGNGGGGCATT	278
DB	241	CGCAGGCCGCGGAGCGCGCCACCTGGNGGGGCATT	278

190 GGAGTGTCTGGCGAGAGTATTGGGACCCATCGGAGCGTAACCTGGATCTCCGAGGCCT 249

```
; PRIOR FILING DATE: 1997-11-07  
; Remaining Prior Application data removed - See File Wrapper or PALM.
```

; PRIOR FILING DATE: 1997-11-07

;
PRIOR

Db 741 GGACACCTGTAGCAGCCTCTTTCACCCATTGCAGAAAGTGCCTGGTTCTCCGGTCCCA 682
Qy 250 GCGGAGCGCCGCCACCTGG 269
Db 681 CTCCTGGGACGCGCAGATGG 662

RESULT 7

US-10-293-244-2233/c
; Sequence 2233, Application US/10293244
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-029
; CURRENT APPLICATION NUMBER: US/10/293,244
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: Not Yet Assigned
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 09/728,422
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 09/693,325
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/663,561
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/654,936
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 09/620,325
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/598,075
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 3960
; SOFTWARE: Custom
; SEQ ID NO 2233
; LENGTH: 1566
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-293-244-2233

Query Match 13.5%; Score 37.6; DB 6; Length 1566;
Best Local Similarity 54.3%; Pred. No. 0.73;
Matches 76; Conservative 0; Mismatches 64; Indels 0; Gaps 0;
Qy 130 GCTCACCTTCTGCCCCCAAGGCGTTGCTGCGGGTGGCCTGCGTGTGCGCTTATGGAG 189
Db 1278 GCCACCTTGAAGGCCACCAGGATGCTCGCTTGGTCTCTCTCCAGGTTGTGGGA 1219
Qy 190 GGAGTGTGTCGCGAGAGTATTGGGACCCCATCGAGCGTAACTGGATCTCCGAGCCT 249
Db 1218 GGACCGTGAGCAGCCTTTGCACCATTTGCAGAAAGTGCCTGGTTCTCCGGTCCCA 1159
Qy 250 GCGGAGCGCGCCACCTGG 269
Db 1158 CTCCTGGGACGCGCAGATGG 1139

RESULT 8

US-10-293-244-2234/c
; Sequence 2234, Application US/10293244
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-029
; CURRENT APPLICATION NUMBER: US/10/293,244
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: Not Yet Assigned
; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: 09/728,422
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 09/693,325
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/663,561
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/654,936
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 09/620,325
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/598,075
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 3960
; SOFTWARE: Custom
; SEQ ID NO 2234
; LENGTH: 1566
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-293-244-2234

Query Match 13.5%; Score 37.6; DB 6; Length 1566;
Best Local Similarity 54.3%; Pred. No. 0.73;
Matches 76; Conservative 0; Mismatches 64; Indels 0; Gaps 0;
Qy 130 GCTCACCTTCTGCCCCCAAGGCGTTGCTGCGGGTGGCCTGCGTGTGCGCTTATGGAG 189
Db 1278 GCCACCTTGAAGGCCACCAGGATGCTCGCTTGGTCTCTCTCCAGGTTGTGGGA 1219
Qy 190 GGAGTGTGTCGCGAGAGTATTGGGACCCCATCGAGCGTAACTGGATCTCCGAGCCT 249
Db 1218 GGACCGTGAGCAGCCTTTGCACCATTTGCAGAAAGTGCCTGGTTCTCCGGTCCCA 1159
Qy 250 GCGGAGCGCGCCACCTGG 269
Db 1158 CTCCTGGGACGCGCAGATGG 1139

RESULT 9

US-10-293-244-265/c
; Sequence 265, Application US/10293244
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang, Y. Tom et al
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-029
; CURRENT APPLICATION NUMBER: US/10/293,244
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: Not Yet Assigned
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 09/728,422
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 09/693,325
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/663,561
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/654,936
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 09/620,325
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/598,075
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 3960
; SOFTWARE: Custom
; SEQ ID NO 265
; LENGTH: 1948

Db 73333 ACCAGGACGAGCGCGTGTGGAGCGTGAGGTGGGACGTGGCGCGGCTCAAGTGGG 73392
QY 191 GAGTGTGTGGCAGAGTATTTCGGGACCCATCGGAGCGTAACCTGGATCTCCGAGGCGCTG 250
Db 73393 CGGAGCCCGGACGCGCCSAGGCGGAGTGCCCAAGGGAGGAGCGCCGAGCTGACCCG 73452
QY 251 GCGGAGCGCGGCCACCTGGNGG 272
Db 73453 GCGACGCGCGGGAGGTTCTGG 73474

RESULT 13
US-10-612-783-2901
; Sequence 2901, Application US/10612783
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; FILE REFERENCE: 38-21(53373)A
; CURRENT APPLICATION NUMBER: US/10/612,783
; CURRENT FILING DATE: 2003-07-02
; NUMBER OF SEQ ID NOS: 7098
; SEQ ID NO 2901
; LENGTH: 955
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4577_91553C.1
US-10-612-783-2901

Query Match 12.6%; Score 35; DB 6; Length 955;
Best Local Similarity 59.0%; Pred. No. 3.2;
Matches 59; Conservative 0; Mismatches 41; Indels 0; Gaps 0;
QY 80 CCCGCGGACCTTCGTTGAGTAACCTGGCGGAGTGGTGGAGCGTGTGCTCACCTTC 139
Db 803 CCGTGGCAGGCGCTTCTTCAACAGGTGCTGGACAGCGTGTGGTCCGATGCTCAAGTTC 862
QY 140 CTGCGCGCAAGGCGTTCCTCGGGTGGCCCTGCGGTGCC 179
Db 863 CTGCAGTCCACGGGTCGCCCGCTCATGCTCACCGGTACC 902

RESULT 14
US-10-612-783-3069
; Sequence 3069, Application US/10612783
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; FILE REFERENCE: 38-21(53373)A
; CURRENT APPLICATION NUMBER: US/10/612,783
; CURRENT FILING DATE: 2003-07-02
; NUMBER OF SEQ ID NOS: 7098
; SEQ ID NO 3069
; LENGTH: 1467
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_12828C.1
US-10-612-783-3069

Query Match 12.6%; Score 35; DB 6; Length 1467;
Best Local Similarity 59.0%; Pred. No. 3.5;
Matches 59; Conservative 0; Mismatches 41; Indels 0; Gaps 0;
QY 80 CCGCGGANACCTTCGTTGAGTAACCTGGCGGAGGTGGTGGAGCGGTGCTCACCTTC 139

Db 502 CCCTCGCAGGCCTTCTTCAACCGTCCTCGACCCGCTGCTGTGGCGCTGCTCAAGTTC 561
QY 140 CTGCGCGCAAGCGCTTGTGCGGGTGGCCCTGCGGTGCC 179
Db 562 TTGCAGTCCACGGGCTCGCCGCTGATGCTCAACGTGTACC 601

RESULT 15
US-10-612-783-3070
; Sequence 3070, Application US/10612783
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; FILE REFERENCE: 38-21(53373)A
; CURRENT APPLICATION NUMBER: US/10/612,783
; CURRENT FILING DATE: 2003-07-02
; NUMBER OF SEQ ID NOS: 7098
; SEQ ID NO 3070
; LENGTH: 2219
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_12829C.1
US-10-612-783-3070

Query Match 12.6%; Score 35; DB 6; Length 2219;
Best Local Similarity 59.0%; Pred. No. 3.7;
Matches 59; Conservative 0; Mismatches 41; Indels 0; Gaps 0;
QY 80 CCCGCGGANACCTTCGTTGAGTAACCTGGCGGAGGTGGTGGAGCGTGTGCTCACCTTC 139
Db 757 CCGTGGCAGGCGCTTCTTCAACCGCTCCCTGGACCGGCTGCTGGTCCCGCTGCTCAAGTTC 816
QY 140 CTGCGCGCAAGCGCTTGTCTCGGGTGGCCCTGCGGTGCC 179
Db 817 TTGCAGTCCACGGGCTCGCCGCTGATGCTCAACGTGTACC 856

Search completed: August 27, 2003, 09:28:25
Job time : 522 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2003, 04:27:30 ; Search time 3740 Seconds
(without alignments)
2435.685 Million cell updates/sec

Title: US-10-042-417A-29

Perfect score: 278
Sequence: 1 ccgtagtactgnttcggc.....cgggccactgnggggcatt 278

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 32997241 seqs, 16383922548 residues

Total number of hits satisfying chosen parameters: 65994482

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_NA_Main.*
1: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq.*
2: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq.old.*
3: /cgn2_6/ptodata/2/pna/US06_COMB.seq.*
4: /cgn2_6/ptodata/2/pna/US07_COMB.seq.*
5: /cgn2_6/ptodata/2/pna/US080_COMB.seq.*
6: /cgn2_6/ptodata/2/pna/US081_COMB.seq.*
7: /cgn2_6/ptodata/2/pna/US082_COMB.seq.*
8: /cgn2_6/ptodata/2/pna/US083_COMB.seq.*
9: /cgn2_6/ptodata/2/pna/US084_COMB.seq.*
10: /cgn2_6/ptodata/2/pna/US085_COMB.seq.*
11: /cgn2_6/ptodata/2/pna/US086_COMB.seq.*
12: /cgn2_6/ptodata/2/pna/US087_COMB.seq.*
13: /cgn2_6/ptodata/2/pna/US088_COMB.seq.*
14: /cgn2_6/ptodata/2/pna/US089_COMB.seq.*
15: /cgn2_6/ptodata/2/pna/US090_COMB.seq.*
16: /cgn2_6/ptodata/2/pna/US091_COMB.seq.*
17: /cgn2_6/ptodata/2/pna/US092A_COMB.seq.*
18: /cgn2_6/ptodata/2/pna/US092B_COMB.seq.*
19: /cgn2_6/ptodata/2/pna/US093A_COMB.seq.*
20: /cgn2_6/ptodata/2/pna/US093B_COMB.seq.*
21: /cgn2_6/ptodata/2/pna/US094_COMB.seq.*
22: /cgn2_6/ptodata/2/pna/US095A_COMB.seq.*
23: /cgn2_6/ptodata/2/pna/US095B_COMB.seq.*
24: /cgn2_6/ptodata/2/pna/US095C_COMB.seq.*
25: /cgn2_6/ptodata/2/pna/US095D_COMB.seq.*
26: /cgn2_6/ptodata/2/pna/US096A_COMB.seq.*
27: /cgn2_6/ptodata/2/pna/US096B_COMB.seq.*
28: /cgn2_6/ptodata/2/pna/US096C_COMB.seq.*
29: /cgn2_6/ptodata/2/pna/US096D_COMB.seq.*
30: /cgn2_6/ptodata/2/pna/US096E_COMB.seq.*
31: /cgn2_6/ptodata/2/pna/US097A_COMB.seq.*
32: /cgn2_6/ptodata/2/pna/US097B_COMB.seq.*
33: /cgn2_6/ptodata/2/pna/US097C_COMB.seq.*
34: /cgn2_6/ptodata/2/pna/US098A_COMB.seq.*
35: /cgn2_6/ptodata/2/pna/US098B_COMB.seq.*
36: /cgn2_6/ptodata/2/pna/US098C_COMB.seq.*
37: /cgn2_6/ptodata/2/pna/US098D_COMB.seq.*
38: /cgn2_6/ptodata/2/pna/US099A_COMB.seq.*
39: /cgn2_6/ptodata/2/pna/US099B_COMB.seq.*
40: /cgn2_6/ptodata/2/pna/US099C_COMB.seq.*
41: /cgn2_6/ptodata/2/pna/US099D_COMB.seq.*
42: /cgn2_6/ptodata/2/pna/US099E_COMB.seq.*
43: /cgn2_6/ptodata/2/pna/US099F_COMB.seq.*

44: /cgn2_6/ptodata/2/pna/US100A_COMB.seq.*
45: /cgn2_6/ptodata/2/pna/US100B_COMB.seq.*
46: /cgn2_6/ptodata/2/pna/US101A_COMB.seq.*
47: /cgn2_6/ptodata/2/pna/US101B_COMB.seq.*
48: /cgn2_6/ptodata/2/pna/US102A_COMB.seq.*
49: /cgn2_6/ptodata/2/pna/US102B_COMB.seq.*
50: /cgn2_6/ptodata/2/pna/US103A_COMB.seq.*
51: /cgn2_6/ptodata/2/pna/US103B_COMB.seq.*
52: /cgn2_6/ptodata/2/pna/US104A_COMB.seq.*
53: /cgn2_6/ptodata/2/pna/US104B_COMB.seq.*
54: /cgn2_6/ptodata/2/pna/US6000_COMB.seq.*
55: /cgn2_6/ptodata/2/pna/US6001_COMB.seq.*
56: /cgn2_6/ptodata/2/pna/US6002_COMB.seq.*
57: /cgn2_6/ptodata/2/pna/US6003_COMB.seq.*
58: /cgn2_6/ptodata/2/pna/US6004_COMB.seq.*
59: /cgn2_6/ptodata/2/pna/US6005_COMB.seq.*
60: /cgn2_6/ptodata/2/pna/US6006_COMB.seq.*
61: /cgn2_6/ptodata/2/pna/US6007_COMB.seq.*
62: /cgn2_6/ptodata/2/pna/US6008_COMB.seq.*
63: /cgn2_6/ptodata/2/pna/US6009_COMB.seq.*
64: /cgn2_6/ptodata/2/pna/US6010_COMB.seq.*
65: /cgn2_6/ptodata/2/pna/US6011_COMB.seq.*
66: /cgn2_6/ptodata/2/pna/US6012_COMB.seq.*
67: /cgn2_6/ptodata/2/pna/US6013_COMB.seq.*
68: /cgn2_6/ptodata/2/pna/US6014_COMB.seq.*
69: /cgn2_6/ptodata/2/pna/US6015_COMB.seq.*
70: /cgn2_6/ptodata/2/pna/US6016_COMB.seq.*
71: /cgn2_6/ptodata/2/pna/US6017_COMB.seq.*
72: /cgn2_6/ptodata/2/pna/US6018_COMB.seq.*
73: /cgn2_6/ptodata/2/pna/US6019_COMB.seq.*
74: /cgn2_6/ptodata/2/pna/US6020_COMB.seq.*
75: /cgn2_6/ptodata/2/pna/US6021_COMB.seq.*
76: /cgn2_6/ptodata/2/pna/US6022_COMB.seq.*
77: /cgn2_6/ptodata/2/pna/US6023A_COMB.seq.*
78: /cgn2_6/ptodata/2/pna/US6023B_COMB.seq.*
79: /cgn2_6/ptodata/2/pna/US6024_COMB.seq.*
80: /cgn2_6/ptodata/2/pna/US6025_COMB.seq.*
81: /cgn2_6/ptodata/2/pna/US6026_COMB.seq.*
82: /cgn2_6/ptodata/2/pna/US6027_COMB.seq.*
83: /cgn2_6/ptodata/2/pna/US6028_COMB.seq.*
84: /cgn2_6/ptodata/2/pna/US6029_COMB.seq.*
85: /cgn2_6/ptodata/2/pna/US6030_COMB.seq.*
86: /cgn2_6/ptodata/2/pna/US6031_COMB.seq.*
87: /cgn2_6/ptodata/2/pna/US6032_COMB.seq.*
88: /cgn2_6/ptodata/2/pna/US6033_COMB.seq.*
89: /cgn2_6/ptodata/2/pna/US6034_COMB.seq.*
90: /cgn2_6/ptodata/2/pna/US6035_COMB.seq.*
91: /cgn2_6/ptodata/2/pna/US6036_COMB.seq.*
92: /cgn2_6/ptodata/2/pna/US6037_COMB.seq.*
93: /cgn2_6/ptodata/2/pna/US6038_COMB.seq.*
94: /cgn2_6/ptodata/2/pna/US6039_COMB.seq.*
95: /cgn2_6/ptodata/2/pna/US6040_COMB.seq.*
96: /cgn2_6/ptodata/2/pna/US6041_COMB.seq.*
97: /cgn2_6/ptodata/2/pna/US6042_COMB.seq.*
98: /cgn2_6/ptodata/2/pna/US6043_COMB.seq.*
99: /cgn2_6/ptodata/2/pna/US6044_COMB.seq.*
100: /cgn2_6/ptodata/2/pna/US6045_COMB.seq.*
101: /cgn2_6/ptodata/2/pna/US6046_COMB.seq.*
102: /cgn2_6/ptodata/2/pna/US6047_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	length	DB ID	Description
1	273	98.2	278	1	PCT-US02-00311-29
2	273	98.2	278	1	PCT-US99-19560-29
3	273	98.2	278	20	PCT-US99-19560-29
4	273	98.2	278	20	US-09-385-219-29

Sequence 29, Appl
Sequence 29, Appl
Sequence 29, Appl


```
RESULT 3
PCT-US99-19560-29
; Sequence 29, Application PC/TUS9919560
; GENERAL INFORMATION:
; APPLICANT: NEW YORK UNIVERSITY
; TITLE OF INVENTION: NOVEL UBIQUITIN LIGASES AS THERAPEUTIC TARGETS
; FILE REFERENCE: 5914-081-228
; CURRENT APPLICATION NUMBER: PCT/US99/19560
; CURRENT FILING DATE: 1999-08-31
; EARLIER APPLICATION NUMBER: 60/098,355
; EARLIER FILING DATE: 1998-08-28
; EARLIER APPLICATION NUMBER: 60/118,568
; EARLIER FILING DATE: 1999-02-03
; EARLIER APPLICATION NUMBER: 60/124,449
; EARLIER FILING DATE: 1999-03-15
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 29
; LENGTH: 278
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: modified_base
; LOCATION: all n positions
; OTHER INFORMATION: n=a, c, g or t
PCT-US99-19560-29

Query Match      98.2%; Score 273; DB 2; Length 278;
Best Local Similarity 100.0%; Pred. No. 5.1e-53;
Matches 278; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCGTAGTACTGGNTTCCGGCGGGCTGGTGAGGAATGGAGCGGTAGNTGCTTGGCGGCG 60
Db 1 CCGTAGTACTGGNTTCCGGCGGGCTGGTGAGGAATGGAGCGGTAGNTGCTTGGCGGCG 60

Qy 61 TCCCGGGNTCCCGTAGACCCCGGGANACCTTCGTGTGAGTAACCTGGCGGAGGTGGT 120
Db 61 TCCCGGGNTCCCGTAGACCCCGGGANACCTTCGTGTGAGTAACCTGGCGGAGGTGGT 120

Qy 121 GGAGCGTGTGCTACCTTCCTCCGCCCAAGCGTTGCTGCGGGTGCGCTGCGTGC 180
Db 121 GGAGCGTGTGCTACCTTCCTCCGCCCAAGCGTTGCTGCGGGTGCGCTGCGTGC 180

Qy 181 CTTATGGAGGAGTGTGTCGCAGAGATTTGCGGACCATCGGAGCGTAACCTGGATCTC 240
Db 181 CTTATGGAGGAGTGTGTCGCAGAGATTTGCGGACCATCGGAGCGTAACCTGGATCTC 240

Qy 241 CGCAGGCTTGGCGGAGCGCGCCACCTGGNGGGGCATT 278
Db 241 CGCAGGCTTGGCGGAGCGCGCCACCTGGNGGGGCATT 278

RESULT 4
US-09-385-219-29
; Sequence 29, Application US/09385219
; GENERAL INFORMATION:
; APPLICANT: Chiau, D.
; APPLICANT: Pagano, M.
; APPLICANT: Latres, E.
; TITLE OF INVENTION: NOVEL UBIQUITIN LIGASES AS THERAPEUTIC TARGETS
; FILE REFERENCE: 5914-081
; CURRENT APPLICATION NUMBER: US/09/385,219A
; CURRENT FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: 60/098,355
; PRIOR FILING DATE: 1998-08-28
; PRIOR APPLICATION NUMBER: 60/118,568
; PRIOR FILING DATE: 1999-02-03
; PRIOR APPLICATION NUMBER: 60/124,449
; PRIOR FILING DATE: 1999-03-15
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 29
; LENGTH: 278
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: modified_base
; LOCATION: all n positions
; OTHER INFORMATION: n=a, c, g or t
US-09-385-219-29

Query Match      98.2%; Score 273; DB 20; Length 278;
Best Local Similarity 100.0%; Pred. No. 5.1e-53;
Matches 278; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCGTAGTA:TCGNTTCCGGCGGGCTGGTGAGGAATGGAGCGGTAGNTGCTTGGCGGCG 60
Db 1 CCGTAGTA:TCGNTTCCGGCGGGCTGGTGAGGAATGGAGCGGTAGNTGCTTGGCGGCG 60

Qy 61 TCCCGGGNTCCCGTAGACCCCGGGANACCTTCGTGTGAGTAACCTGGCGGAGGTGGT 120
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 29
```

```
; LENGTH: 278
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: all n positions
; OTHER INFORMATION: n=a, c, g or t
US-09-385-219-29

Query Match      98.2%; Score 273; DB 20; Length 278;
Best Local Similarity 100.0%; Pred. No. 5.1e-53;
Matches 278; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCGTAGTACTGGNTTCCGGCGGGCTGGTGAGGAATGGAGCGGTAGNTGCTTGGCGGCG 60
Db 1 CCGTAGTACTGGNTTCCGGCGGGCTGGTGAGGAATGGAGCGGTAGNTGCTTGGCGGCG 60

Qy 61 TCCCGGGNTCCCGTAGACCCCGGGANACCTTCGTGTGAGTAACCTGGCGGAGGTGGT 120
Db 61 TCCCGGGNTCCCGTAGACCCCGGGANACCTTCGTGTGAGTAACCTGGCGGAGGTGGT 120

Qy 121 GGAGCGTGTGCTACCTTCCTCCGCCCAAGCGTTGCTGCGGGTGCGCTGCGTGC 180
Db 121 GGAGCGTGTGCTACCTTCCTCCGCCCAAGCGTTGCTGCGGGTGCGCTGCGTGC 180

Qy 181 CTTATGGAGGAGTGTGTCGCAGAGATTTGCGGACCATCGGAGCGTAACCTGGATCTC 240
Db 181 CTTATGGAGGAGTGTGTCGCAGAGATTTGCGGACCATCGGAGCGTAACCTGGATCTC 240

Qy 241 CGCAGGCTTGGCGGAGCGCGCCACCTGGNGGGGCATT 278
Db 241 CGCAGGCTTGGCGGAGCGCGCCACCTGGNGGGGCATT 278

RESULT 5
US-09-385-219A-29
; Sequence 29, Application US/09385219A
; GENERAL INFORMATION:
; APPLICANT: Chiau, D.
; APPLICANT: Pagano, M.
; APPLICANT: Latres, E.
; TITLE OF INVENTION: NOVEL UBIQUITIN LIGASES AS THERAPEUTIC TARGETS
; FILE REFERENCE: 5914-081
; CURRENT APPLICATION NUMBER: US/09/385,219A
; CURRENT FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: 60/098,355
; PRIOR FILING DATE: 1998-08-28
; PRIOR APPLICATION NUMBER: 60/118,568
; PRIOR FILING DATE: 1999-02-03
; PRIOR APPLICATION NUMBER: 60/124,449
; PRIOR FILING DATE: 1999-03-15
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 29
; LENGTH: 278
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: modified_base
; LOCATION: all n positions
; OTHER INFORMATION: n=a, c, g or t
US-09-385-219A-29

Query Match      98.2%; Score 273; DB 20; Length 278;
Best Local Similarity 100.0%; Pred. No. 5.1e-53;
Matches 278; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCGTAGTA:TCGNTTCCGGCGGGCTGGTGAGGAATGGAGCGGTAGNTGCTTGGCGGCG 60
Db 1 CCGTAGTA:TCGNTTCCGGCGGGCTGGTGAGGAATGGAGCGGTAGNTGCTTGGCGGCG 60

Qy 61 TCCCGGGNTCCCGTAGACCCCGGGANACCTTCGTGTGAGTAACCTGGCGGAGGTGGT 120
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 29
```

Db 61 TCCGGGNTCCCTCCGTAGACCCCGGGANACCTTCGTGTTGAGTAACCTGCGGAGGTGGT 120
QY 121 GGAGCGTGTCTCACCTTCCTCCCGCCAAAGCGCTTGCTGCGGGTGCCCTGCGTGCCG 180
Db 121 GGAGCGTGTCTCACCTTCCTCCCGCCAAAGCGCTTGCTGCGGGTGCCCTGCGTGCCG 180
QY 181 CTTATGAGGAGGTGTGTCGCGAGAGTATTGCGGACCCATCGGAGCGTAACCTGGATCTC 240
Db 181 CTTATGAGGAGGTGTGTCGCGAGAGTATTGCGGACCCATCGGAGCGTAACCTGGATCTC 240
QY 241 CGCAGGCTGCGGAGCGCGCCACCTGGNGGGGCATT 278
Db 241 CGCAGGCTGCGGAGCGCGCCACCTGGNGGGGCATT 278

RESULT 6

US-10-042-417-29
; Sequence 29, Application US/10042417
; GENERAL INFORMATION:
; APPLICANT: Pagano, M.
; TITLE OF INVENTION: METHODS TO IDENTIFY COMPOUNDS USEFUL FOR THE TREATMENT OF
; FILE OF INVENTION: PROLIFERATIVE AND DIFFERENTIATIVE DISORDERS
; FILE REFERENCE: 5914-090-999
; CURRENT APPLICATION NUMBER: US/10/042,417
; PRIOR FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: 60/260,179
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 29
; LENGTH: 278
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: modified_base
; LOCATION: all n positions
; OTHER INFORMATION: n=a, c, g or t
US-10-042-417-29

Query Match 98.2%; Score 273; DB 44; Length 278;
Best Local Similarity 100.0%; Pred. No. 5.1e-53;
Matches 278; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGTAGTACTGNTTCCGGCGGGTGTGAGGAATGGAGCGGTAGNTGCTTGGCGCGAG 60
Db 1 CCGTAGTACTGNTTCCGGCGGGTGTGAGGAATGGAGCGGTAGNTGCTTGGCGCGAG 60
QY 61 TCCGGGNTCCCTCCGTAGACCCCGGGANACCTTCGTGTTGAGTAACCTGCGGAGGTGGT 120
Db 61 TCCGGGNTCCCTCCGTAGACCCCGGGANACCTTCGTGTTGAGTAACCTGCGGAGGTGGT 120
QY 121 GGAGCGTGTCTCACCTTCCTCCCGCCAAAGCGCTTGCTGCGGGTGCCCTGCGTGCCG 180
Db 121 GGAGCGTGTCTCACCTTCCTCCCGCCAAAGCGCTTGCTGCGGGTGCCCTGCGTGCCG 180
QY 181 CTTATGAGGAGGTGTGTCGCGAGAGTATTGCGGACCCATCGGAGCGTAACCTGGATCTC 240
Db 181 CTTATGAGGAGGTGTGTCGCGAGAGTATTGCGGACCCATCGGAGCGTAACCTGGATCTC 240
QY 241 CGCAGGCTGCGGAGCGCGCCACCTGGNGGGGCATT 278
Db 241 CGCAGGCTGCGGAGCGCGCCACCTGGNGGGGCATT 278

RESULT 7

US-09-644-871-7991
; Sequence 7991, Application US/09644871
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Pan, Yang
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: 1600.1167-001

; CURRENT APPLICATION NUMBER: US/09/644,871
; CURRENT FILING DATE: 2000-08-28
; PRIOR APPLICATION NUMBER: 60/151,059
; PRIOR FILING DATE: 1999-08-27
; NUMBER OF SEQ ID NOS: 9739
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7991
; LENGTH: 708
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; LOCATION: (1)...(708)
; OTHER INFORMATION: n = A,T,C or G
US-09-644-871-7991

Query Match 84.5%; Score 235; DB 28; Length 708;
Best Local Similarity 93.9%; Pred. No. 3e-44;
Matches 263; Conservative 0; Mismatches 15; Indels 2; Gaps 2;
QY 1 CCGTAGTACTGNTTCCGGCGGGTGTGAGGAATGGAGCGGTAGNTGCTTGGCGCGAG 60
Db 26 CCGTAGGAGTACTGTTCCGGCGGGTGTGAGGAATGGAGCGGTAGNTGCTTGGCGCGAG 85
QY 61 TCCCG-GGNTCTCCGTAGACCCCGCGA-NACCTTCGTGTTGAGTAACCTGCGGAGGTG 118
Db 86 TCGCGCGCTCCTCCGTAGACCCCGCGAGCACTTCGTGTTGAGTAACCTGCGGAGGTG 145
QY 119 GTGAGCGTGTCTCACCTTCCTGCGCCGCAAGGCGTGTGCTGCGGGTGCCCTGCGTGTC 178
Db 146 GTGAGCGTGTCTCACCTTCCTGCGCCGCAAGGCGTGTGCTGCGGGTGCCCTGCGTGTC 205
QY 179 CGCTTATGAGGAGTGTGTGCGCAGAGTATTGCGGACCCATCGGAGCGTAACCTGGATC 238
Db 206 CGCTTATGAGGAGTGTGTGCGCAGAGTATTGCGGACCCATCGGAGCGTAACCTGGATC 265
QY 239 TCCGAGGCGTGGCGGAGCGCGCCACCTGGNGGGGCATT 278
Db 266 TCCGAGGCGTGGCGGAGCGCGCCACCTGGAGGGGCATT 305

RESULT 8

US-09-652-124-8036
; Sequence 8036, Application US/09652124
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Welch, Nadine S.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: 1600.1179-001
; CURRENT APPLICATION NUMBER: US/09/652,124
; CURRENT FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/151,131
; PRIOR FILING DATE: 1999-08-30
; NUMBER OF SEQ ID NOS: 9868
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8036
; LENGTH: 708
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; LOCATION: (1)...(708)
; OTHER INFORMATION: n = A,T,C or G
US-09-652-124-8036

Query Match 84.5%; Score 235; DB 28; Length 708;
Best Local Similarity 93.9%; Pred. No. 3e-44;
Matches 263; Conservative 0; Mismatches 15; Indels 2; Gaps 2;
QY 1 CCGTAGTACTGNTTCCGGCGGGTGTGAGGAATGGAGCGGTAGNTGCTTGGCGCGAG 60
Db 26 CCGTAGGAGTACTGTTCCGGCGGGTGTGAGGAATGGAGCGGTAGNTGCTTGGCGCGAG 85

RESULT 10
US-09-652-355-8346
; Sequence 8346. Application US/09652355
; GENERAL INFORMATION:
; APPLICANT: Shyjan, Andrew W.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES

```

RESULT 11
US-08-276-163A-1116
; Sequence 1116, Application US/08276163A
; GENERAL INFORMATION:
; APPLICANT: Haseltine, William
; APPLICANT: Fitzgerald, Lisa
; APPLICANT: Adams, Mark
; APPLICANT: Lee, Normal
; APPLICANT: Fuldner, Rebecca
; APPLICANT: Fleischmann, Robert
; APPLICANT: Bult, Carol
; APPLICANT: Blake, Judy
; APPLICANT: White, Owen
; APPLICANT: Clayton, Rebecca
; APPLICANT: Pellegriano, Susan
; TITLE OF INVENTION: Human Genes, Sequences, and Expression
; NUMBER OF SEQUENCES: 1532
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carulla, Byrne, Bain, Gilfillan, Cecchi, Stewart, &
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 5.0
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:

```

```
; APPLICATION NUMBER: US/08/276.163A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Olstein, Elliot
; REGISTRATION NUMBER: 24,025
; REFERENCE/DOCKET NUMBER: 325800-175
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (201) 994-1700
; TELEFAX: (201) 994-1744
; INFORMATION FOR SEQ ID NO: 1116:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 391 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-276-163A-1116

Query Match      83.5%; Score 232; DB 7; Length 391;
Best Local Similarity 97.2%; Pred. No. 1.5e-43;
Matches 244; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

Qy 28 TGAGGAATGAGCGCGGTAGTTCCTGCGCGGAGTCCCGGNTCTCCGTAGACCCGCGGA 87
Db 1 TGAGGAATGAGCGCGGTAGTTCCTGCGCGGAGTCCCGGNTCTCCGTAGACCCGCGGA 59

Qy 88 NACCTTCGTGTGAGTAACCTGCGCGGAGTGGTGGAGCGTGTCTACCTTCCTGCGCAGT 147
Db 60 GACCTTCGTGTGAGTAACCTGCGCGGAGTGGTGGAGCGTGTCTACCTTCCTGCGCAGT 119

Qy 148 CAAGCGCTTGTGCGCGGTGCGCTGTCGCGTTCATGGAGGAGTGTCTGCGCAGT 207
Db 120 CAAGCGCTTGTGCGCGGTGCGCTGTCGCGTTCATGGAGGAGTGTCTGCGCAGT 179

Qy 208 ATTGCGGACCCATCGGAGCGTAACCTGGATCTCCGAGGCGCTGGCGAGGCCGCCACCT 267
Db 180 ATTGCGGACCCATCGGAGCGTAACCTGGATCTCCGAGGCGCTGGCGAGGCCGCCACCT 239

Qy 268 GNGGGGCATT 278
Db 240 GGAGGGGNATT 250

RESULT 12
US-08-276-163B-1116
; Sequence 1116, Application US/08276163B
; GENERAL INFORMATION:
; APPLICANT: Adams, Mark D.
; APPLICANT: Bult, Carol J.
; APPLICANT: Blake, Judith A.
; APPLICANT: Fitzgerald, Lisa
; APPLICANT: Fleischmann, Robert D.
; APPLICANT: Fraser, Claire M.
; APPLICANT: Fuldner, Rebecca A.
; APPLICANT: Gocayne, Jeannine D.
; APPLICANT: Kelley, Jenny
; APPLICANT: Kirkness, Ewen F.
; APPLICANT: Sutton, Granger G., III
; APPLICANT: Haseltine, William A.
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steve
; APPLICANT: Dillon, Patrick J.
; APPLICANT: Li, Haodong
; TITLE OF INVENTION: Human Genes, Sequences and Expression Products
; NUMBER OF SEQUENCES: 15314
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
```

```
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 Inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/276.163B
; FILING DATE: July 15, 1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Michele M. Wales
; REGISTRATION NUMBER: P-43,975
; REFERENCE/DOCKET NUMBER: PO-14
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 1116:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 391 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-276-163B-1116

Query Match      83.5%; Score 232; DB 7; Length 391;
Best Local Similarity 97.2%; Pred. No. 1.5e-43;
Matches 244; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

Qy 28 TGAGGAATGAGCGCGGTAGTTCCTGCGCGGAGTCCCGGNTCTCCGTAGACCCGCGGA 87
Db 1 TGAGGAATGAGCGCGGTAGTTCCTGCGCGGAGTCCCGGNTCTCCGTAGACCCGCGGA 59

Qy 88 NACCTTCGTGTGAGTAACCTGCGCGGAGTGGTGGAGCGTGTCTACCTTCCTGCGCAGT 147
Db 60 GACCTTCGTGTGAGTAACCTGCGCGGAGTGGTGGAGCGTGTCTACCTTCCTGCGCAGT 119

Qy 148 CAAGCGCTTGTGCGCGGTGCGCTGTCGCGTTCATGGAGGAGTGTCTGCGCAGT 207
Db 120 CAAGCGCTTGTGCGCGGTGCGCTGTCGCGTTCATGGAGGAGTGTCTGCGCAGT 179

Qy 208 ATTGCGGACCCATCGGAGCGTAACCTGGATCTCCGAGGCGCTGGCGAGGCCGCCACCT 267
Db 180 ATTGCGGACCCATCGGAGCGTAACCTGGATCTCCGAGGCGCTGGCGAGGCCGCCACCT 239

Qy 268 GNGGGGCATT 278
Db 240 GGAGGGGNATT 250

RESULT 13
US-08-276-163D-1116
; Sequence 1116, Application US/08276163D
; GENERAL INFORMATION:
; APPLICANT: Adams, et. al.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products
; FILE REFERENCE: PO14
; CURRENT APPLICATION NUMBER: US/08/276.163D
; CURRENT FILING DATE: 1994-07-15
; NUMBER OF SEQ ID NOS: 15314
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1116
; LENGTH: 391
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (20)
; OTHER INFORMATION: n equals a,t,g, or c
```

NAME/KEY: misc feature
LOCATION: (37)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (167)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (247)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (282)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (353)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (360)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (367)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (384)
OTHER INFORMATION: n equals a,t,g, or c
US-08-276-163D-1116

Query Match 83.5%; Score 232; DB 7; Length 391;
Best Local Similarity 97.2%; Pred. No. 1.5e-43;
Matches 244; Conservative 0; Mismatches 6; Indels 1; Gaps 1;
QY 28 TGAGGAATGAGCGCGTAGNTGCTTGGCGGAGTCCCGGGNTCCCGTAGACCCGCGGA 87
DB 1 TGAGGAATGAGCGCGTAGNTGC-TGGCGGAGTCCCGGGCTCCCGTAGACCCGCGGA 59
QY 88 NACCTTCGTGTGAGTAACCTGGCGGAGGTGGTGGAGCGTGTGCTACCTTCCTGCCCGC 147
DB 60 GACCTTCGTGTGAGTAACCTGGCGGAGGTGGTGGAGCGTGTGCTACCTTCCTGCCCGC 119
QY 148 CAAGCGCTGTGCGGGTGGCGCTCGTGTGCCGCTTATGGAGGAGTGTGTGCCAGAGT 207
DB 120 CAAGCGCTGTGCGGGTGGCGCTCGTGTGCCGCTTATGGAGGAGTGTGTGCCAGAGT 179
QY 208 ATTCCGACCCATCGGAGCGTAACCTGGATCTCCGAGGCGCTGGCGAGGCGCGCCACCT 267
DB 180 ATTCCGACCCATCGGAGCGTAACCTGGATCTCCGAGGCGCTGGCGAGGCGCGCCACCT 239
QY 268 GNGGGGCGATT 278
DB 240 GGAGGGGNATT 250

RESULT 14
US-09-840-145-1116
Sequence 1116, Application US/09840145
GENERAL INFORMATION:
APPLICANT: Adams, et. al.
TITLE OF INVENTION: Human Genes, Sequences, and Expression Products
FILE REFERENCE: P014C1
CURRENT APPLICATION NUMBER: US/09/840,145
PRIOR FILING DATE: 2001-04-24
PRIOR APPLICATION NUMBER: 08/276,163
PRIOR FILING DATE: 1994-07-15
NUMBER OF SEQ ID NOS: 15314
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1116
LENGTH: 391
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (20)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature

LOCATION: (37)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (167)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (247)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (282)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (353)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (360)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (367)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc_feature
LOCATION: (384)
OTHER INFORMATION: n equals a,t,g, or c
US-09-840-145-1116

Query Match 83.5%; Score 232; DB 35; Length 391;
Best Local Similarity 97.2%; Pred. No. 1.5e-43;
Matches 244; Conservative 0; Mismatches 6; Indels 1; Gaps 1;
QY 28 TGAGGAATGAGCGCGTAGNTGCTTGGCGGAGTCCCGGGNTCCCGTAGACCCGCGGA 87
DB 1 TGAGGAATGAGCGCGTAGNTGC-TGGCGGAGTCCCGGGCTCCCGTAGACCCGCGGA 59
QY 88 NACCTTCGTGTGAGTAACCTGGCGGAGGTGGTGGAGCGTGTGCTACCTTCCTGCCCGC 147
DB 60 GACCTTCGTGTGAGTAACCTGGCGGAGGTGGTGGAGCGTGTGCTACCTTCCTGCCCGC 119
QY 148 CAAGCGCTGTGCGGGTGGCGCTCGTGTGCCGCTTATGGAGGAGTGTGTGCCAGAGT 207
DB 120 CAAGCGCTGTGCGGGTGGCGCTCGTGTGCCGCTTATGGAGGAGTGTGTGCCAGAGT 179
QY 208 ATTCCGACCCATCGGAGCGTAACCTGGATCTCCGAGGCGCTGGCGAGGCGCGCCACCT 267
DB 180 ATTCCGACCCATCGGAGCGTAACCTGGATCTCCGAGGCGCTGGCGAGGCGCGCCACCT 239
QY 268 GNGGGGCGATT 278
DB 240 GGAGGGGNATT 250

RESULT 15
US-09-465-877-10811
Sequence 10811, Application US/09465877
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
FILE REFERENCE: 20411-781
CURRENT APPLICATION NUMBER: US/09/465,877
CURRENT FILING DATE: 1999-12-16
NUMBER OF SEQ ID NOS: 15920
SOFTWARE: Hy-patent.pl Version 3.1
SEQ ID NO 10811
LENGTH: 451
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(451)
OTHER INFORMATION: n = A,T,C or G
US-09-465-877-10811

Query Match 83.5%; Score 232; DB 21; Length 451;

Best Local Similarity 95.5%; Pred. No. 1.5e-43;			
Matches	257;	Conservative	0;
	Mismatches	10;	Indels
		2;	Gaps
			2;
Qy	12	GNTTCCGGCGGCTGGTGAGGAATGGAGCCGGTAGNTGCTTGGCGCGAGTCCCG-GGNTC	70
Db	37		
Qy	71	CTCCGTAGACCCGGCGGA-NACCTTCGTGTTGAGTAACCTGGCGGAGGTGGGAGCGTGT	129
Db	97	CTCCGTAGACCCGGCGGACCTTCGTGTTGAGTAACCTGGCGGAGGTGGGAGCGTGT	156
Qy	130	GCTCACCTTCTGCCCCCGCAAGCGTTGCTGCCGGTGGCCCTGCGCTTATGGAG	189
Db	157	GCTCACCTTCTGCCCCCGCAAGCGTTGCTGCCGGTGGCCCTGCGCTTATGGAG	216
Qy	190	GGAGTGTGTGGCGAGAGTATTGCCGACCCATCGGAGCGTAACCTGGATCTCCGAGGCCT	249
Db	217	GGAGTGTGTGGCGAGAGTATTGCCGACCCATCGGAGCGTAACCTGGATCTCCGAGGCCT	276
Qy	250	GGCGGAGCGCGCCACCTGGNGGGGCATT	278
Db	277	GGCGGAGCGCGCCACCTGGAGGGGCATT	305

Search completed: August 27, 2003, 09:19:32
Job time : 3744 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2003, 06:41:34 ; Search time 210 Seconds
(without alignments)
3037.151 Million cell updates/sec

Title: US-10-042-417A-29

Perfect score: 278

Sequence: 1 ccgtagtactgnttcggc.....cggccactgngggcgatt 278

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1533700 seqs, 1147125425 residues

Total number of hits satisfying chosen parameters: 3067400

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	273	98.2	278	13	US-10-042-417-29
2	229	82.4	1256	14	US-10-153-668-159
3	229	82.4	1418	14	US-10-153-668-161
4	113.6	40.9	2086	11	US-09-974-879-89
5	113.6	40.9	2086	11	US-09-305-736-89
6	40.6	14.6	9025608	14	US-10-156-761-1
7	37.8	13.6	1725	14	US-10-156-761-2436
8	37.6	13.5	698	10	US-09-764-868-250
9	37.6	13.5	698	10	US-09-764-868-562
10	37.4	13.5	2385	14	US-10-156-761-6087
11	37.4	13.5	9025608	14	US-10-156-761-1
12	36.6	13.2	474	11	US-09-918-995-7426
13	36.6	13.2	1095	14	US-10-156-761-1222
14	36.2	13.0	1173	14	US-10-156-761-7021
15	35.2	12.7	2394	14	US-10-156-761-5612
16	35	12.6	1242	14	US-10-166-087-33

c	17	35	12.6	32539	14	US-10-166-087-1	Sequence 1, Appli
	18	34.2	12.3	898	12	US-10-017-161-661	Sequence 661, App
	19	34.2	12.3	1275	14	US-10-156-761-3923	Sequence 3923, Ap
c	20	33.8	12.2	827	13	US-10-027-632-167037	Sequence 167037,
	21	33.8	12.2	827	13	US-10-027-632-167038	Sequence 167038,
c	22	33.6	12.1	720	14	US-10-156-761-6622	Sequence 6622, Ap
	23	33.6	12.1	1620	9	US-09-815-242-4084	Sequence 4084, Ap
	24	33	11.9	3552	14	US-10-156-761-4242	Sequence 4242, Ap
c	25	32.8	11.8	1302	12	US-10-309-175-3	Sequence 3, Appli
	26	32.8	11.8	23238	14	US-10-156-761-415	Sequence 415, App
	27	32.8	11.8	100000	14	US-10-156-761-15103	Sequence 15103, A
	28	32.6	11.7	898	12	US-10-017-161-681	Sequence 681, App
	29	32.6	11.7	3528	14	US-10-156-761-3536	Sequence 3536, Ap
	30	32.4	11.7	1023	14	US-10-156-761-3452	Sequence 3452, Ap
	31	32.4	11.7	3066	13	US-10-027-632-112941	Sequence 112941,
	32	32.4	11.7	3731	9	US-09-822-288A-1	Sequence 1, Appli
	33	32.4	11.7	48667	9	US-09-822-288A-3	Sequence 3, Appli
	34	32.2	11.6	903	14	US-10-156-761-1980	Sequence 1980, Ap
c	35	32.2	11.6	1350	14	US-10-156-761-7401	Sequence 7401, Ap
	36	32.2	11.6	1640	12	US-10-237-852-6	Sequence 6, Appli
	37	32.2	11.6	2469	14	US-10-274-878-1	Sequence 1, Appli
	38	32	11.5	621	9	US-09-815-242-3982	Sequence 3982, Ap
	39	32	11.5	624	9	US-09-815-242-7555	Sequence 7555, Ap
	40	32	11.5	3068	14	US-10-027-828-16	Sequence 16, Appl
c	41	32	11.5	5970	13	US-10-108-605-210	Sequence 210, App
	42	32	11.5	6282	13	US-10-108-605-212	Sequence 212, App
c	43	32	11.5	11238	14	US-10-205-032-15	Sequence 15, Appl
	44	32	11.5	60196	14	US-10-205-032-1	Sequence 1, Appli
	45	31.8	11.4	290	9	US-09-294-093B-1299	Sequence 1299, Ap

ALIGNMENTS

RESULT 1

US-10-042-417-29
; Sequence 29, Application US/10042417
; Publication No. US20020123082A1
; GENERAL INFORMATION:
; APPLICANT: Pagano, M.
; TITLE OF INVENTION: METHODS TO IDENTIFY COMPOUNDS USEFUL FOR THE TREATMENT OF
; TITLE OF INVENTION: PROLIFERATIVE AND DIFFERENTIATIVE DISORDERS
; FILE REFERENCE: 5914-090-999
; CURRENT APPLICATION NUMBER: US/10/042,417
; CURRENT FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: 60/260,179
; PRIOR FILING DATE: 2001-01-5
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 29
; LENGTH: 278
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: all n positions
; OTHER INFORMATION: n=a, c, g or t
US-10-042-417-29

Query Match 98.2%; Score 273; DB 13; Length 278;
Best Local Similarity 100.0%; Pred. No. 2.8e-69;
Matches 278; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CCGTAGTACTCGNTTCCGGCGGCTGGTGAGCAATGGAGCGGTAGTTCCTTGGCGCGAG 60
Db 1 CCGTAGTACTCGNTTCCGGCGGCTGGTGAGCAATGGAGCGGTAGTTCCTTGGCGCGAG 60
Qy 61 TCCCGGGNCTCCGTAGACCCCGGANACCTTCGTGTGAGTAACCTGGCGGAGGTGCT 120
Db 61 TCCCGGGNCTCCGTAGACCCCGGANACCTTCGTGTGAGTAACCTGGCGGAGGTGCT 120
Qy 121 GGAGCGTGTGCTACCTTCCTCCGCCCAAGCGGTGCTGCGGGTGGCCTCGTGTGCCG 180
Db 121 GGAGCGTGTGCTACCTTCCTCCGCCCAAGCGGTGCTGCGGGTGGCCTCGTGTGCCG 180

Db 121 GGAGCGTGTCTACCTTCCTGCCCGCAAGCGTGTGCTGCGGGTGGCGCTTGCCTGTGCGG 180
Qy 181 CTTATGAGGAGTGTGCGCAGAGTATTGCGGACCCATCGGAGCGTAACCTGGATCTC 240
Db 181 CTTATGAGGAGTGTGCGCAGAGTATTGCGGACCCATCGGAGCGTAACCTGGATCTC 240
Qy 241 CGAGCGCTGGGGAGGCGCGCACCTCGNGGGGCATT 278
Db 241 CGAGCGCTGGGGAGGCGCGCACCTCGNGGGGCATT 278

RESULT 2

US-10-153-668-159
; Sequence 159, Application US/10153668
; Publication No. US20030092616A1
; GENERAL INFORMATION:
; APPLICANT: HONDA, Goichi
; APPLICANT: MATSUDA, Akio
; APPLICANT: MURAMATSU, Shuji
; APPLICANT: ISHIZAWA, Kenya
; TITLE OF INVENTION: STAT6 Activating Gene
; FILE REFERENCE: 1254-0207P
; CURRENT APPLICATION NUMBER: US/10/153,668
; PRIOR FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 60/293,172
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 60/316,031
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/328,403
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: JP 2001-157043
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: JP 2001-260681
; PRIOR FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: JP 2001-313175
; NUMBER OF SEQ ID NOS: 488
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 159
; LENGTH: 1256
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (19)..(1227)

US-10-153-668-159

Query Match 82.4%; Score 229; DB 14; Length 1256;
Best Local Similarity 95.8%; Pred. No. 1.2e-56;
Matches 254; Conservative 0; Mismatches 9; Indels 2; Gaps 2;

Qy 16 CCGCGCGGCTGGTGAGGAATGGAGCCGTAGTCTTTCGCGCGAGTCCCG-GGNTCTCC 74
Db 1 CCGCGCGGCTGGTGAGGAATGGAGCCGTAGTCTTTCGCGCGAGTCCCGCGGCTCC 60
Qy 75 GTAGACCCGCGGA-NACCTTCGTGTTGAGTAACCTGCGGAGGTGGTGGAGCGTGTGCTC 133
Db 61 GTAGACCCGCGGACACCTTCGTGTTGAGTAACCTGCGGAGGTGGTGGAGCGTGTGCTC 120
Qy 134 ACCTTCCTGCCCGCAAGCGGTGCTGCGGGTGGCCCTGCGGTGTCCTTATGAGGGAG 193
Db 121 ACCTTCCTGCCCGCAAGCGGTGCTGCGGGTGGCCCTGCGGTGTCCTTATGAGGGAG 180
Qy 194 TGTGTGCGCAGAGTATTGCGGACCCATCGGAGCGTAACCTGGATCTCTCCGAGCGCTGGCG 253
Db 181 TGTGTGCGCAGAGTATTGCGGACCCATCGGAGCGTAACCTGGATCTCTCCGAGCGCTGGCG 240
Qy 254 GAGCGCGGCCACCTCGNGGGGCATT 278
Db 241 GAGCGCGGCCACCTGGAGGGGCATT 265

RESULT 3

US-10-153-668-161
; Sequence 161, Application US/10153668
; Publication No. US20030092616A1
; GENERAL INFORMATION:
; APPLICANT: HONDA, Goichi
; APPLICANT: MATSUDA, Akio
; APPLICANT: MURAMATSU, Shuji
; APPLICANT: ISHIZAWA, Kenya
; TITLE OF INVENTION: STAT6 Activating Gene
; FILE REFERENCE: 1254-0207P
; CURRENT APPLICATION NUMBER: US/10/153,668
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 60/293,172
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 60/316,031
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/328,403
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: JP 2001-157043
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: JP 2001-260681
; PRIOR FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: JP 2001-313175
; NUMBER OF SEQ ID NOS: 488
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 161
; LENGTH: 1418
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (19)..(846)

US-10-153-668-161

Query Match 82.4%; Score 229; DB 14; Length 1418;
Best Local Similarity 95.8%; Pred. No. 1.2e-56;
Matches 254; Conservative 0; Mismatches 9; Indels 2; Gaps 2;

Qy 16 CCGCGCGGCTGGTGAGGAATGGAGCCGTAGTCTTTCGCGCGAGTCCCG-GGNTCTCC 74
Db 1 CCGCGCGGCTGGTGAGGAATGGAGCCGTAGTCTTTCGCGCGAGTCCCGCGGCTCC 60
Qy 75 GTAGACCCGCGGA-NACCTTCGTGTTGAGTAACCTGCGGAGGTGGTGGAGCGTGTGCTC 133
Db 61 GTAGACCCGCGGACACCTTCGTGTTGAGTAACCTGCGGAGGTGGTGGAGCGTGTGCTC 120
Qy 134 ACCTTCCTGCCCGCAAGCGGTGCTGCGGGTGGCCCTGCGGTGTCCTTATGAGGGAG 193
Db 121 ACCTTCCTGCCCGCAAGCGGTGCTGCGGGTGGCCCTGCGGTGTCCTTATGAGGGAG 180
Qy 194 TGTGTGCGCAGAGTATTGCGGACCCATCGGAGCGTAACCTGGATCTCTCCGAGCGCTGGCG 253
Db 181 TGTGTGCGCAGAGTATTGCGGACCCATCGGAGCGTAACCTGGATCTCTCCGAGCGCTGGCG 240
Qy 254 GAGCGCGGCCACCTCGNGGGGCATT 278
Db 241 GAGCGCGGCCACCTGGAGGGGCATT 265

RESULT 4

US-09-974-879-89
; Sequence 89, Application US/09974879
; Publication No. US20030028003A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 125 Human Secreted Proteins
; FILE REFERENCE: P2020P2
; CURRENT APPLICATION NUMBER: US/09/974,879
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/239,893
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 09/818,663

; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (4187715)
; OTHER INFORMATION: a, t, c, g, other or unknown
US-10-156-761-1

Query Match 14.6%; Score 40.6; DB 14; Length 9025608;
Best Local Similarity 50.3%; Pred. No. 0.015;
Matches 97; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

Qy 65 GGNVCTCCGTAGACCGCGGANACCTTCGTGTGAGTAACCTGGCGGAGGTGGTGAG 124
Db 2985290 GTGCTCATAGCTCTCTCGCATCGCTACGTGCGGTGCCACCGGTGCGTGTGAACGAG 2985349

Qy 125 CGTGTGCTCACTTCCTGCCCGCAAGCGTTGCTGCGGTGGCGCTCGTGTGCCGTTA 184
Db 2985350 TGTGGGCACTCTTGGCGCACTCGCGGTCTTGGCGGAAACCGCGAGTCCGTCCATCG 2985409

Qy 185 TGGAGGAGTGTGTGCGCAGAGTATTGCGGACCCATCGGAGCGTAACTTGGATCTCCGCA 244
Db 2985410 CGGACCGTGTTCGTCGCGGCGAGCGCGGTGCCCTGGCTGACGCTCTCCCTCTCGGG 2985469

Qy 245 GGCCTGGCGGAGG 257
Db 2985470 CGACCGTCTGGG 2985482

RESULT 7
US-10-156-761-2436/c
; Sequence 2436, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 2436
; LENGTH: 1725
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1725)
US-10-156-761-2436

Query Match 13.6%; Score 37.8; DB 14; Length 1725;
Best Local Similarity 50.0%; Pred. No. 0.1;
Matches 93; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

Qy 72 TCGTAGACCGCGGANACCTTCTGTGAGTAACCTGGCGGAGGTGGTGGAGCGTGTGC 131
Db 1725 TAGCTCTCCCTGGCATCGCTAGCTGCGGTGCCACCGGTGCTGTGAACGAGTGTGTGC 1666

Qy 132 TCACCTTCCTGCCCGCAAGCGGTGCTGCGGTGGCGCTGCGTGTGCCCTTATGAGGG 191
Db 1665 ACCCTCTGCGCACTCCGGGTCTTGGCGGAAACCGCGAGTCCGTCCATCGCGACCG 1606

Qy 192 AGTGTGTGCGCAGAGTATTGCGGACCCATCGGAGCGTAACTGGATCTCCGAGCGCTGG 251

Db 1605 TGTTCCTGGCGCGCAGCCCGGTGCTCGGTGAGCGCTGCTCTGGGGCGGACCGT 1546
Qy 252 CGGAGG 257
Db 1545 CTTGGG 1540

RESULT 8
US-09-764-868-250
; Sequence 250, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: ROSEN ET AL.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT232
; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 250
; LENGTH: 698
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (616)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-868-250

Query Match 13.5%; Score 37.6; DB 10; Length 698;
Best Local Similarity 51.6%; Pred. No. 0.12;
Matches 82; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

Qy 14 TTCCGGCGGCTGTGTGAGGAATGAGCGGTAGTGTGCGGCGAGTCCCGGGNTCCTC 73
Db 227 TCCAGCGCGCGCGGCAATTTCTGATCGGTGATGAGCGGTGCTGCCGCTACGTG 286

Qy 74 CGTAGACCGCGGANACCTTCGTGTGAGTAACCTGCGGAGGTGGTGGAGCGTGTGCTC 133
Db 287 TGCACGACAGAGTCTCACCAGCTGTTCAGCGCGGTCAAGCGGTGCTGTCCGGGTAC 346

Qy 134 ACCTTCCTGCCCGCCAAAGCGTTGTGCGGGTGGCCCTGC 172
Db 347 AGCTACTTCCCAACAGCGGTGTAACCTCGGTGCGCTCC 385

RESULT 9
US-09-764-868-562
; Sequence 562, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: ROSEN ET AL.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT232
; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 562
; LENGTH: 698
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (600)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (611)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (615)

; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (644)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-868-562

Query Match 13.5%; Score 37.6; DB 10; Length 698;

Best Local Similarity 51.6%; Pred. No. 0.12;

Matches 82; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 14 TTCGCGGGCTGGTAGGAATGCGGCGGTGCTTCCGCGGAGTCCGCGGNTCCCTC 73

Db 227 TCCAGCGCCCGGGCATTTCTCGATGCGGTGATGAGCGGGTGTGCGCGCTAGGTG 286

QY 74 CGTAGACCCCGCGGANACCTTCGTGTTGAGTAACTGGCGGAGGTGGTGGAGCGGTGTGCTC 133

Db 287 TGCAAGCAACAGCATCTCACCGAGCTGTTGAGCGCGGTCAAGCGGTGTGTCGGGTAC 346

QY 134 ACCTTCTGCGCGCAAGCGGTGCTGCGGGTGGCGCTGC 172

Db 347 AGCTACTTCCCAACAGCGGTTGAACCTCGTGGCTCC 385

RESULT 10

US-10-156-761-6087

; Sequence 6087, Application US/10156761

; Publication No. US20030119018A1

; GENERAL INFORMATION:

; APPLICANT: OMURA, SATOSHI

; APPLICANT: IKEDA, HARUO

; APPLICANT: ISHIKAWA, JUN

; APPLICANT: HORIKAWA, HIROSHI

; APPLICANT: SHIBA, TADAYOSHI

; APPLICANT: SAKAKI, YOSHIYUKI

; APPLICANT: HATTORI, MASAHIRA

; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

; FILE REFERENCE: 249-262

; CURRENT APPLICATION NUMBER: US/10/156,761

; CURRENT FILING DATE: 2002-05-29

; PRIOR APPLICATION NUMBER: JP 2001-204089

; PRIOR FILING DATE: 2001-05-30

; PRIOR APPLICATION NUMBER: JP 2001-272697

; PRIOR FILING DATE: 2001-08-02

; NUMBER OF SEQ ID NOS: 15109

; SEQ ID NO 6087

; LENGTH: 2385

; TYPE: DNA

; ORGANISM: Streptomyces avermitilis

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)..(2385)

US-10-156-761-6087

Query Match

Best Local Similarity 13.5%; Score 37.4; DB 14; Length 2385;

Matches 71; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 53 GCGCGAGTCCCGGNTCCCTCGTAGACCGCGGANACCTTCGTGTTGAGTAACTCGCG 112

Db 1822 GTGGCGGCTCGGCGAGCGCGGTCCGCCCTGTGTACGGGATCTCTGTGGACCTGGTG 1881

QY 113 GAGGTGGTGGAGCGGTGTGCTACCTTCCTGCCGCCCAAGCGGTTGCTGCGGGTGGCCTGC 172

Db 1882 GGGCTCGTGTGCTGTGGTGTGCTGCCCTTCCTGCCGATGTGTGGCGGATGCGGGTCCGGGCC 1941

QY 173 GTGTGCCGC 181

Db 1942 GTACGGCTC 1950

RESULT 11

US-10-156-761-1/c

; Sequence 1, Application US/10156761

; Publication No. US20030119018A1

; GENERAL INFORMATION:

; APPLICANT: OMURA, SATOSHI

; APPLICANT: IKEDA, HARUO

; APPLICANT: ISHIKAWA, JUN

; APPLICANT: HORIKAWA, HIROSHI

; APPLICANT: SHIBA, TADAYOSHI

; APPLICANT: SAKAKI, YOSHIYUKI

; APPLICANT: HATTORI, MASAHIRA

; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

; FILE REFERENCE: 249-262

; CURRENT APPLICATION NUMBER: US/10/156,761

; CURRENT FILING DATE: 2002-05-29

; PRIOR APPLICATION NUMBER: JP 2001-204089

; PRIOR FILING DATE: 2001-05-30

; PRIOR APPLICATION NUMBER: JP 2001-272697

; PRIOR FILING DATE: 2001-08-02

; NUMBER OF SEQ ID NOS: 15109

; SEQ ID NO 1

; LENGTH: 9025608

; TYPE: DNA

; ORGANISM: Streptomyces avermitilis

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (4187715)

; OTHER INFORMATION: a, t, c, g, other or unknown

US-10-156-761-1

Query Match

Best Local Similarity 13.5%; Score 37.4; DB 14; Length 9025608;

Matches 71; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 53 GCGCGAGTCCCGGNTCCCTCGTAGACCGCGGANACCTTCGTGTTGAGTAACTCGCG 112

Db 7347813 GTGGCGGCTCGGCGAGCGCGGTCCGCCCTGTGTACGGGATCTCTGTGGACCTGGTG 7347754

QY 113 GAGGTGGTGGAGCGGTGTGCTACCTTCCTGCCGCCCAAGCGGTTGCTGCGGGTGGCCTGC 172

Db 7347753 GGGCTCGTGTGCTGTGCTGCTGCCCTTCCTGCCGATGTGTGGCGGATGCGGGTCCGGGCC 7347694

QY 173 GTGTGCCGC 181

Db 7347693 GTACGGCTC 7347685

RESULT 12

US-09-918-995-7426/c

; Sequence 7426, Application US/09918995

; Publication No. US20030073623A1

; GENERAL INFORMATION:

; APPLICANT: Hyseq, Inc.

; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED

; FILE REFERENCE: 20411-756

; CURRENT APPLICATION NUMBER: US/09/918,995

; CURRENT FILING DATE: 2001-07-30

; PRIOR APPLICATION NUMBER: US/09/235,076

; PRIOR FILING DATE: 1999-01-20

; NUMBER OF SEQ ID NOS: 38054

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 7426

; LENGTH: 474

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (1)..(474)

; OTHER INFORMATION: n = A,T,C or G

US-09-918-995-7426

Query Match

Best Local Similarity 13.2%; Score 36.6; DB 11; Length 474;

Matches 111; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

;	APPLICANT:	HORIKAWA, HIROSHI
;	APPLICANT:	SHIBA, TADAYOSHI
;	APPLICANT:	SAKAKI, YOSHIOYUKI
;	APPLICANT:	HATTORI, MASAHIRA
;	TITLE OF INVENTION:	NOVEL POLYNUCLEOTIDES
;	FILE REFERENCE:	249-262
;	CURRENT APPLICATION NUMBER:	US/10/156,761
;	CURRENT FILING DATE:	2002-05-29
;	PRIOR APPLICATION NUMBER:	JP 2001-204089
;	PRIOR FILING DATE:	2001-05-30
;	PRIOR APPLICATION NUMBER:	JP 2001-272697
;	PRIOR FILING DATE:	2001-08-02
;	NUMBER OF SEQ ID NOS:	15109
;	SEQ ID NO 7021	
;	LENGTH:	1173
;	TYPE:	DNA
;	ORGANISM:	Streptomyces avermitilis
;	FEATURE:	
;	NAME/KEY:	CDS
;	LOCATION:	(1)..(1173)
;	US-10-156-761-7021	

Query Match	13.0%;	Score 36.2;	DB 14;	Length 1173;
Best Local Similarity	45.6%;	Pred. No. 0.29;		
Matches 119;	Conservative 0;	Mismatches 142;	Indels 0;	Gaps 0;
Oy	16	CCGCGGGCTGGTGCAGGAATGGAGCCGTAGTNTCCTTGCGCGGAGTCGCCGGNTCCTCGG	75	
Db	765	CTGCGCGGGCCGACCCGCTGGTGGGGCTGCCCATCACGCGGGGATCGCGCTGGTCT	824	
Oy	76	TAGACC CGCGGANACCTTCGTTGAGTAACCTGGCGGAGTGGTGGAGCGCTGCTCCAC	135	
Db	825	CGCGGACGCGCGCGAGGTGTTCCGCGGGTGATGACSCCTCGACCCGCGCTGGT	884	
Oy	136	CTTCTCGCGCCAAAGGCTTGCTGCGGGTGGCTGGGTGCGCGCTTATCGAGGGAGTG	195	
Db	885	GGACGCGGGCGAGCGGCACTGGAGGAGTGTCCCGGGCTACGCGACGCTCGGTGAGCTGG	944	
Oy	196	TGTCGCGAGAGTATFTGCGGACCCATCGAGGGTAACCTGGATCTCCGAGCCCTGGCGGA	255	
Db	945	GTTGCGGTGATCGGGCACCGGCTGCGGGCGGAGGTGCGCGCTCGTCTCGACGCGGAGGT	1004	
Oy	256	GGCGGGCCACTGGNGGGCA	276	
Db	1005	GAGCGTGGCCAGCGCATCA	1025	

RESULT 15

US-10-156-761-5612/c

; Sequence 5612, Application US/10156761

; Publication No. US20030119018A1

; GENERAL INFORMATION:

; APPLICANT: OMURA, SATOSHI

; APPLICANT: IKEDA, HARUO

; APPLICANT: ISHIKAWA, JUN

; APPLICANT: HORIKAWA, HIROSHI

; APPLICANT: SHIBA, TADAYOSHI

; APPLICANT: SAKAKI, YOSHIOYUKI

; APPLICANT: HATTORI, MASAHIRA

; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

; FILE REFERENCE: 249-262

; CURRENT APPLICATION NUMBER: US/10/156,761

; CURRENT FILING DATE: 2002-05-29

; PRIOR APPLICATION NUMBER: JP 2001-204089

; PRIOR FILING DATE: 2001-05-30

; PRIOR APPLICATION NUMBER: JP 2001-272697

; PRIOR FILING DATE: 2001-08-02

; NUMBER OF SEQ ID NOS: 15109

; SEQ ID NO 5612

; LENGTH: 2394

; TYPE: DNA

; ORGANISM: Streptomyces avermitilis

; FEATURE:

```

; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1173)
US-10-156-761-7021

Query Match 13.0%; Score 36.2; DB 14; Length 1173;
Best Local Similarity 45.6%; Pred. No. 0.29;
Matches 119; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

QY 16 CCGCGCGGCTGGTCAGGAATGGAGCCGCTAGTGTCTTGGCGCGAGTCCCGGGNTCCTCGC 75
DB 765 CTGCGCGCGCGCCACCCGCTGGTGGGGCTGGCCATCAGCGCGGGATCGCGCTGGTGCT 824

QY 76 TAGACCCGCGGANACCTTCGTGTTGAGTAACTTGGCGGAGGTGGTGGAGCGTGTGCTCAC 135
DB 825 GCGGAGCGCGCGCGAGGTGTTCCGCGGGGTGATGCACCCCTCGACCCGCGCTGCT 884

QY 136 CTTCCTCCCGCCAAGCGTGTCTGCGGTGGCCTGGCTGTGCCGTATATCGAGGGAGTG 195
DB 885 GGACCGGGCCGAGCGGCGACTGGAGGAGGTGCCCGCGGTACCGGACGTGGGTGAGCTCG 944

QY 196 TGTGCGCAGAGTATTGGCGGACCCCATCGGAGCGTAACCTGGAGTCTCCGCGAGCGCTGGCGGA 255
DB 945 GTTCGCTGGATCGGCGCACCGGCTGCGGGCGGAGGTGGCGCTCGTCTCGACGCGGAGGT 1004

QY 256 GCGCGCCACCTGGNGGGCA 276
DB 1005 GAGCGTGGCCAGCGGCATCA 1025

RESULT 15
US-10-156-761-5612/c
; Sequence 5612, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 5612
; LENGTH: 2394
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:

```

```

RESULT 15
US-10-156-761-5612/c
; Sequence 5612, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 5612
; LENGTH: 2394
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:

```

```

: PRIOR FILING DATE: 2001-05-30
: PRIOR APPLICATION NUMBER: JP 2001-272697
: PRIOR FILING DATE: 2001-08-02
: NUMBER OF SEQ ID NOS: 15109
: SEQ ID NO 5612
: LENGTH: 2394
: TYPE: DNA
: ORGANISM: Streptomyces avermitilis
: FEATURE:

```

; NAME/KEY: CDS
; LOCATION: (1)..(2394)
US-10-156-761-5612

Query Match	12.7%	Score 35.2;	DB 14;	Length 2394;
Best Local Similarity	51.6%	Pred. No. 0.57;	74;	Indels 0; Gaps 0;
Matches 79;	Conservative 0;	Mismatches	74;	

QY	80	CCCGCGGANACCTTCGTGTTGAGTAACCTGGCGGAGGTGGTGAGCGTGTGCTCACCTTC	139
Db	2267	CCGGAGCGCGGTTTCAGGTGGAGTACGTCCCGCAGGTCGTGCGATCCGGCGGTCAACGTC	2208
QY	140	CTGCCCGCCCAAGGCGTTGCTCGCGGTGGCTGCGGTGGCCCTTATGAGGAGGTGTGTG	199
Db	2207	AGCTTGTGCTACTGGGTGCTGGTGTGCTCCCGCGTGTGCGATGTCAGGTAGAAGGTC	2148
QY	200	CGCAGAGTATTGCGGACCCATCGGAGCGTAACC	232
Db	2147	AGGCTGGCCTTGCAAGCGGCGGATGTCACC	2115

Search completed: August 27, 2003, 09:33:32
Job time : 250 secs

THIS PAGE BLANK (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2003, 01:24:48 : Search time 314 Seconds.
(without alignments)

390.779 Million cell updates/sec

Title: US-10-042-417A-29

Perfect score: 278

Sequence: 1 cgtagtactggttcggc.....cgggccacctgnggggcatt 278

Scoring table: IDENTITY_NUC

Gapex 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents.NA.*

- 1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
- 2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
- 3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
- 4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
- 5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq.*
- 6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	41.4	14.9	4403765	3	US-09-103-840A-2
C 2	41.4	14.9	4411529	3	US-09-103-840A-1
C 3	38.6	13.9	1788	4	US-09-252-991A-15365
C 4	38.6	13.9	1383	4	US-09-252-991A-15469
C 5	38.6	13.9	1503	4	US-09-252-991A-15409
C 6	35.6	12.8	1014	4	US-09-252-991A-15280
C 7	35.6	12.8	1842	4	US-09-252-991A-15508
C 8	35.6	12.8	43280	2	US-08-804-227C-1
C 9	34.8	12.5	534	4	US-09-252-991A-15379
C 10	34.8	12.5	4403765	3	US-09-103-840A-2
C 11	34.8	12.5	4411529	3	US-09-103-840A-1
C 12	34.2	12.3	471	4	US-09-252-991A-2383
C 13	34.2	12.3	480	4	US-09-252-991A-2536
C 14	34.2	12.3	882	4	US-09-252-991A-2222
C 15	34.2	12.3	984	4	US-09-252-991A-2298
C 16	33.4	12.0	465	4	US-09-252-991A-8014
C 17	33.4	12.0	1029	4	US-09-252-991A-8069
C 18	33.4	12.0	1167	4	US-09-252-991A-8122
C 19	33.4	12.0	1194	4	US-09-252-991A-8118
C 20	33.4	12.0	1221	4	US-09-252-991A-7020
C 21	33.4	12.0	1845	4	US-09-252-991A-8187
C 22	33.2	11.9	474	4	US-09-252-991A-6996
C 23	33.2	11.9	744	4	US-09-252-991A-13502
C 24	33.2	11.9	849	4	US-09-252-991A-13952
C 25	33	11.9	285	4	US-09-252-991A-11646
C 26	33	11.9	570	4	US-09-252-991A-2398
C 27	33	11.9	696	4	US-09-252-991A-12736

```

28 33 11.9 708 4 US-09-252-991A-13051 Sequence 13051, A
29 33 11.9 729 4 US-09-252-991A-11790 Sequence 11790, A
30 33 11.9 864 4 US-09-252-991A-12904 Sequence 12904, A
31 33 11.9 999 4 US-09-252-991A-2600 Sequence 2600, Ap
32 33 11.9 1047 4 US-09-252-991A-2516 Sequence 2516, Ap
33 33 11.9 1095 4 US-09-252-991A-2312 Sequence 2312, Ap
34 33 11.9 1260 4 US-09-252-991A-11855 Sequence 11855, A
35 33 11.9 1986 4 US-09-252-991A-12512 Sequence 12512, A
36 33 11.9 2085 4 US-09-252-991A-12962 Sequence 12962, A
37 33 11.9 2211 4 US-09-252-991A-6288 Sequence 6288, Ap
38 33 11.9 2328 4 US-09-252-991A-6020 Sequence 6020, Ap
39 32.6 11.7 1113 4 US-09-252-991A-2947 Sequence 2947, Ap
40 32.6 11.7 1158 4 US-09-252-991A-5439 Sequence 5439, Ap
41 32.6 11.7 1386 4 US-09-252-991A-2736 Sequence 2736, Ap
42 32.6 11.7 1548 4 US-09-252-991A-5452 Sequence 5452, Ap
43 32.6 11.7 1641 4 US-09-252-991A-5477 Sequence 5477, Ap
44 32.6 11.7 3459 4 US-09-016-434-1363 Sequence 1363, Ap
45 32.2 11.6 1405 1 US-08-390-162-3 Sequence 3, Appli

```

ALIGNMENTS

```

RESULT 1
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24356-20007.00
; CURRENT APPLICATION NUMBER: US/09/103.840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Query Match 14.9%; Score 41.4; DB 3; Length 4403765;
Best Local Similarity 48.6%; Pred. No. 0.1;
Matches 108; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

QY 53 GCGCGAGTCCGGGNTCTCGGTAGACCCGCGGANACCTTCGTGTGAGTAACCTGGCG 112
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2709675 GCTTCCGTTTGGACGTACGCTGTGTCGCCGCGCGTGTGATGACGCGC 172
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 113 GAGTGTGTGAGCGTGTGCTCACCTTCTCCGCCCAAGCGTGTGTCGGGTGCGCTGC 172
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2709615 GCGGGGATGTGTGGCCGATCGGCAACGTGTGTGCTGTGCTGCGTGTGCGGCTC 2709556
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 173 GTGTGCGCTTATGTGAGGAGTGTGTGCGCAGAGTATTCGGAGCCATCGGAGCGTAACC 232
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2709555 GCGCGCGGCGACACTGTGTGTGTGTGCGCGCGCGGTGTGTCGCGTGTGCGGACTGGT 2709496
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 233 TGGATCTCCGAGGCGCTGGCGGAGCGCGCCACCTTGGNGGGG 274
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2709495 TCGATCAGCGCGCGCTGTGTCGCGTGTGTCGCGTGTGTCGCGTGTGTCGCGCGC 2709454
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 2
US-09-103-840A-1/c
; Sequence 1, Application US/09103840A

```

Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Query Match 14.9%; Score 41.4; DB 3; Length 4411529;
Best Local Similarity 48.6%; Pred. No. 0.1; Indels 0; Gaps 0;
Matches 108; Conservative 0; Mismatches 114

Qy 53 GCGCGAGTCCCGGNTCTCCGTAGACCGCGGANACCTTCGTTGAGTAACCTGGCG 112
Db 2712862 GCGTCCGTTTGGACGTACGCTGCTCCCGGCGCTGGTACGTGATTGTACCGCG 2712803

Qy 113 GAGTGTGTGAGCGTGTCTACCTTCTGCGCGCAAGCGTGTGCTGCGGTGGCGCTGC 172
Db 2712802 GCGGGATCTGTGCGCGCATCGCAACGTGTGTGCTGCTGCGTGTGGCGCTC 2712743

Qy 173 GTGTGCGCTTATGAGGAGGTGTGCGCAGAGTATTGCGGACCCATCGAGGTAACC 232
Db 2712742 GCGCGCGCGCACTGTGTGTGTGCGCGCGCGTGTGCGACGCTCCGCGACTGGT 2712683

Qy 233 TGGATCTCCGAGCGCTGCGGAGGCGCGCCACCTGGNGGG 274
Db 2712682 TCGATCAGCGCGCGCTGCTGCGGTATGTGTGGCGCG 2712641

RESULT 3
US-09-252-991A-15365/c
; Sequence 15365, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 15365
; LENGTH: 768
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-15365

Query Match 13.9%; Score 38.6; DB 4; Length 768;
Best Local Similarity 49.7%; Pred. No. 0.14; Indels 0; Gaps 0;
Matches 95; Conservative 0; Mismatches 96

Qy 82 CCGCGANACCTTCGTGTGAGTAACCTGCGGAGGTGGTGGAGCGTGTGCTACCTTCCT 141
Db 628 CCGCGCCATCGTCGGTGTGAGTCCGCGCCCATGTGTGCTGCGCGCTGGGCTTCCT 569

Qy 142 GCCCGCAAGCGTGTGCTGCGGTGCGCTGCGCTTATGAGGAGGTGTGTGCG 201

568 GATCACCGCGTGGGCTGCGGTGATCACCGTATGATCCCTGATCGCCCAAGGTGCGCGGTTTC 509

Qy 202 CAGAGTATTGGGACCCATCGGACGTAACCTGGATCTCCGACGCGCTGGCGGAGGCGGG 261

Db 508 GGTGACGCGCTTCAGCCATCCGATCGGAGGTATGCGGCGGCTGCTGGCGGCGGTCTG 449

Qy 262 CCACCTGGNGG 272

Db 448 CTACCTGGCGG 438

RESULT 4
US-09-252-991A-15469
; Sequence 15469, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 15469
; LENGTH: 1383
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-15469

Query Match 13.9%; Score 38.6; DB 4; Length 1383;
Best Local Similarity 49.7%; Pred. No. 0.15; Indels 0; Gaps 0;
Matches 95; Conservative 0; Mismatches 96

Qy 82 CCGCGANACCTTCGTGTGAGTAACCTGCGGAGGTGGTGGAGCGTGTGCTACCTTCCT 141
Db 153 CCGCGCCATCGTCGGTGTGAGTCCGCGCCCATGTGTGCTGCGGCGCTGGGCTTCCT 212

Qy 142 GCCCGCAAGCGTGTGCTGCGGTGCGCTGCGCTTATGAGGAGGTGTGTGCG 201

Db 213 GATCACGCGCTGGGCTGCGGTGATCAGCTGATCGCCCTGCGCAAGGTGCGCGGTTTC 272

Qy 202 CAGAGTATTGGGACCCATCGGAGCGTAACCTGGATCTCCGACGCGCTGGCGGAGGCGGG 261

Db 273 GGTGACGCGCTTCAGCCATCCGATCGGAGGTATGCGGCGGCTGCTGGCGGCGGTCTG 332

Qy 262 CCACCTGGNGG 272

Db 333 CTACCTGGCGG 343

RESULT 5
US-09-252-991A-15409
; Sequence 15409, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 15409
; LENGTH: 1503
; TYPE: DNA

```
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-15409

Query Match      13.9%; Score 38.6; DB 4; Length 1503;
Best Local Similarity 49.7%; Pred. No. 0.15; Mismatches 0; Gaps 0;
Matches 95; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

Qy 82 CGCGGACCTTCGTTGAGTAACCTGGCGGAGGTGGTGAGCGTGTGCTCACTTCCT 141
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 115 CCCGCCCATCTCGCTTTCAGTCCGCCCCCATGTGTGGCTGGCGGCTGGGCTTCCT 174
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 142 GCCCGCAAGCGTGTGCTGGGGTGGCTGGGTGTGGCGCTTATGGAGGAGTGTGGCG 201
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 175 GATCACCGCGTGGGGCTGGCGGTGATCACCGTGTATGCCCTGCCAAGGTTCGGCGGTT 234
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 202 CAGAGTATTGGGACCATCGAGCGTAACCTTGGATCTCCGACGCTGGCGGCGCGG 261
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 235 GGTGACGCCCTCAGCCATCGATCGGAGTATGCCGGCGGCTGTGGCGCGGTTG 294
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 262 CCACCTGGNGG 272
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 295 CTACCTGGCGG 305
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 6
US-09-252-991A-15280/c
; Sequence 15280, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 15280
; LENGTH: 1014
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-15280

Query Match      12.8%; Score 35.6; DB 4; Length 1014;
Best Local Similarity 47.4%; Pred. No. 0.87; Mismatches 112; Indels 0; Gaps 0;
Matches 101; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

Qy 16 CCGCGGGCTGGTGAGGAATGAGCGGTAGTGTCTTCCGCGGAGTCCCGGGNTCCCTCG 75
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 656 CTCTGCTGTTATGCCAGGCGACCTGTCAGCGCTTGGAGGAGCGGCGCGCTGGG 597
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 76 TAGACCCCGCGGANACCTTCGTGTTAGTAACCTGGCGGAGGTGGTGAGCGTGTGCTCAC 135
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 596 TAGCAGCGCGGAGTGGCGCTGTCGTCAGCACAACTGACCCAGGTGGCGGCGCGCTT 537
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 136 CTTCCTCCCGCAAGCGTGTGCTGGGTGGCTGGCTGTGCGCTTATGGAGGAGTG 195
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 536 GTTCTGGCGGCCCGCGCTGTCGAAGTGGCGGACCTGCGCGCCCTCGGAGAGCGTTG 477
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 196 TGTGCGCAGATATTGGGACCCCATCGGAGCGT 228
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 476 TGGGGGCTGCGCCACGGGCGCTGGAGGCGT 444
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 7
US-09-252-991A-15508
; Sequence 15508, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
```

```
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 15508
; LENGTH: 1842
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-15508

Query Match      12.8%; Score 35.6; DB 4; Length 1842;
Best Local Similarity 47.4%; Pred. No. 0.96;
Matches 101; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

Qy 16 CCGCGGGCTGGTGAGGAATGAGCGGTAGTGTCTTCCGCGGAGTCCCGGGNTCCCTCG 75
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 477 CTCTGCTGTTATGCCAGGCGACCTGTCAGCGCTTGGAGGAGCGGCGCGCTGGG 536
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 76 TAGACCCCGCGGANACCTTCGTGTTAGTAACCTGGCGGAGGTGGTGAGCGTGTGCTCAC 135
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 537 TAGCAGCGCGGAGTGGCGCTGTCGTCAGCACAACTGACCCAGGTGGCGGCGCGCTT 596
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 136 CTTCCTCCCGCAAGCGTGTGCTGGGTGGCTGGCTGTGCGCTTATGGAGGAGTG 195
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 597 GTTCTGGCGGCCCGCGCTGTCGAAGTGGCGGACCTGCGCGCCCTCGGAGAGCGTTG 656
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 196 TGTGCGCA3AGTATTGGCGGACCCATCGGAGCGT 228
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 657 TGGCGGGCTGCGCCACGGGCGCTGGAGGCGT 689
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 8
US-08-804-227C-1
; Sequence 1, Application US/08804227C
; Patent No. 5876991
; GENERAL INFORMATION:
; APPLICANT: DeHoff, Bradley S.
; APPLICANT: Kuhstoss, Stuart A.
; APPLICANT: Rosteck, Paul R., Jr.
; APPLICANT: Sutton, Kimberly L.
; TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: THOMAS G. PLANT 1501
; STREET: LILLY CORPORATE CENTER
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: USA
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: ASCII(DOS) text only
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804,227C
; FILING DATE: February 21, 1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Plant, Thomas, G.
; REGISTRATION NUMBER: 35,784
; REFERENCE/DOCKET NUMBER: X-8231
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-2459
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 43280 base-pairs
```

```
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 816..14234
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 14351..19945
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 20010..31199
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 31232..36067
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 36249..41774
; US-08-804-227C-1

Query Match      12.8%; Score 35.6; DB 2; Length 43280;
Best Local Similarity 50.3%; Pred. No. 1.6;
Matches 86; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

QY 106 CTTGGCGGAGTGTGGAGCGTGTCTACCTTCTGCCGCCCAAGCGGTTGCTGCCGGT 165
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 35347 CTTGGTGGCGGATGCGACTGGAGCGGTTGCCGCCGCCCTTACCCGGTTCCGGCCCG 35406

QY 166 GGCCTGCGTGTGCCGCTATATGGAGGAGTGTGCGGAGAGTATTGGGACCCATCGGAG 225
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 35407 TGGCGTGTATCTCCCGCGTGTGCTCGGACGCGCGGGGAGGCGGAGCGCCCGGGA 35466

QY 226 CGTAACCTGTATCTCCGCGAGCGCTGGGGAGGCGCGCACCTGNGGGGCA 276
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 35467 CGGCACGCTGTTCGCCGCCGGGTTCCGGCGCGCGGCCACTGAGCGGCA 35517

RESULT 9
US-09-252-991A-15379/c
; Sequence 15379, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 15379
; LENGTH: 534
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-15379

Query Match      12.58; Score 34.8; DB 4; Length 534;
Best Local Similarity 48.2%; Pred. No. 1.3;
Matches 93; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 36 GGAGCCGGTAGTGTTCGCGCGAGTCCCGGGNTCCCTAGACCCGCGGANACCTTCG 95
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 523 GCAGCCTGCAGCCCTGGGAGGAGGCGGCGCGCTGGGTAGCAGCGCGGAGCTGGCGC 464

QY 96 TGTGTAGTAACCTGGCGGAGGTGGTGAGCGTGTGCTACCTTCTCCGCGCCCAAGCGT 155
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 463 TGGTCCAGCACAACTGACCCAGTGGCGCGAGCCCTTGTGCTGGCGCCCGCCGCGC 404

QY 156 TCCTGCGGGTGGCGCTGCGTGTGCCGCTTATGGAGGAGTGTGTGCGCAGAGTATTCGGA 215
```

```
||||| 403 TGCTCGAAGTGGCGGACCTCGCGGCCCTCGGAGAGCGTTGTCCGGGGCTGGCCACGGGGC 344
QY 216 CCCATCGGAGCGT 228
    |||||
Db 343 GCCTGGAGGGCGT 331

RESULT 10
US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Query Match      12.58; Score 34.8; DB 3; Length 4403765;
Best Local Similarity 51.7%; Pred. No. 5.3;
Matches 75; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 45 AGNTGCTTGGCGGAGTCCCGGGNTCTCCGTAGACCCGCGGANACCTTCTGTGTGAGTA 104
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4185051 ATGCTTTGGCCGGGTGACGCGGTAGCGCGGCTATCATTTGGGCATGGTTA 4185110

QY 105 ACCTGGCGGAGTGTGGAGCGTGTGCTCACCTTCTCCCGCCCAAGCGTGTGCTGCGGG 164
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4185111 CCGCGCGCGGCTGTGGGCCATGCCGCGCACCACCGTGACCGCGCTGCGCGCGTGC 4185170

QY 165 TGGCTGCGTGTGCCGCTTATGGAG 189
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4185171 TGATCACCATCCGACGCTCATGGTG 4185195

RESULT 11
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1
```



```
Query Match      12.5%; Score 34.8; DB 3; Length 4411529;
Best Local Similarity 51.7%; Pred. No. 5.3;
Matches 75; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 45 AGNTGCTTGGCGGAGTCCCGGNTCTCGTAGACCGGGANACCTTCGTGTGAGTA 104
DB 4192804 ATGCTTTTGGCCGGTGACGGGTAGCGCGATCGCGGCTATCATTTGGGCATGTTA 4192863
QY 105 ACCTGGCGGAGGTGGTGGAGCTGTGCTACCTTCCTGCCCGCAAGCGCTTGTGCGGG 164
DB 4192864 CCGCGCGCGGCTGTGGCCATGCGCGCACCGAGTACCGCGTCCGCGCGTCCGGA 4192923
QY 165 TGGCCTGCGTGTGCCGCTTATGGAG 189
DB 4192924 TGATCACCATCCGAGCTCATGGTG 4192948

RESULT 12
US-09-252-991A-2383/c
; Sequence 2383, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 2383
; LENGTH: 471
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-2383

Query Match      12.3%; Score 34.2; DB 4; Length 471;
Best Local Similarity 48.4%; Pred. No. 1.8;
Matches 90; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 19 GCGGGCTGTGAGGAATGGACCGGTAGTCTTCCGGCGAGTCCCGGNTCTCCGTAG 78
DB 338 GCGGGCTGTTCGGTAGTCGCGGGGAGGTTCGCACCTTGTTCACGATCCAGGGCAG 279
QY 79 ACCCGCGGANACCTTCGTGTGAGTAACCTGCGCGGAGGTGTGGAGCGTGTGCACCTT 138
DB 278 GCCACGCTGCTCGGGCGGTTTCGAGGATCGCGCGGGTGGCGGCTGAAGGTGCGCGAG 219
QY 139 CTGCGCGCAAGCGGTGTGCTGGGTTGCGGTGCTGCGCTTATGAGGAGTGTGT 198
DB 218 CAGATCGCCACTTCCTGGCAGTGTGGCGGCTTCGGCTGCGCGTGGCGCGCGCAG 159
QY 199 GCGCAG 204
DB 158 GCGCGG 153

RESULT 13
US-09-252-991A-2536
; Sequence 2536, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
```

```
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 2536
; LENGTH: 480
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-2536

Query Match      12.3%; Score 34.2; DB 4; Length 480;
Best Local Similarity 48.4%; Pred. No. 1.8;
Matches 90; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 19 GCGGGCTGTGAGGAATGGACCGGTAGTCTTCCGGCGAGTCCCGGNTCTCCGTAG 78
DB 182 GCGGGCTGTTCGGTAGTCGCGGGGAGGTTCGCACCTTGTTCACGATCCAGGGCAG 241
QY 79 ACCCGCGGANACCTTCGTGTGAGTAACCTGCGCGGAGGTGTGGAGCGTGTGCACCTT 138
DB 242 GCCACGGTCTGCTCGGGCGTTCGAGGATCGCCCGCGGTGCGGGCTGAAGGTGCGCCAG 301
QY 139 CTGCGCGCAAGCGGTGTGCTGGGTTGCGGTGCGCTTATGAGGAGTGTGT 198
DB 302 CAGATCGTCCACTTCCTGGCAGTGTGGCGGTTTCGGCTGCGCTGCGCGCGCGCAG 361
QY 199 GCGCAG 304
DB 362 GCGCGG 367

RESULT 14
US-09-252-991A-2222/c
; Sequence 2222, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 2222
; LENGTH: 882
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-2222

Query Match      12.3%; Score 34.2; DB 4; Length 882;
Best Local Similarity 48.4%; Pred. No. 2;
Matches 90; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 19 GCGGGCTGTGAGGAATGGACCGGTAGTCTTCCGGCGAGTCCCGGNTCTCCGTAG 78
DB 490 GCGGGCTGTTCGGTAGTCGCGGGGAGGTTCGCACCTTGTTCACGATCCAGGGCAG 431
QY 79 ACCCGCGGANACCTTCGTGTGAGTAACCTGCGCGGAGGTGTGGAGCGTGTGCACCTT 138
DB 430 GCCACGGTCTGCTCGGGCGTTCGAGGATCGCCCGCGGTGCGGGCTGAAGGTGCGCCAG 371
QY 139 CTGCGCGCAAGCGGTGTGCTGGGTTGCGGTGCGCTTATGAGGAGTGTGT 198
DB 370 CAGATCGCCACTTCCTGGCAGTGTGGCGGTTTCGGCTGCGCTGCGCGCGCGCAG 311
QY 199 GCGCAG 204
DB 310 GCGCGG 305
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2003, 00:14:19 ; Search time 1460 Seconds
(without alignments)
514.003 Million cell updates/sec

Title: US-10-042-417a-29
Perfect score: 278
Sequence: 1 ccgtactgctgntccggc.....cgccacctgnggggcatt 278

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_19Jun03.*
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	273	98.2	278	21	AAZ93360
2	273	98.2	278	24	AAH41051
3	232	83.5	1298	22	AAH08680
4	232	83.5	1298	24	ABN59644
5	232	83.5	1306	22	ABA08916
6	232	83.5	1306	22	AAK52841
7	232	83.5	1418	25	ABZ24703
8	232	83.5	1486	22	AAH16815

Sequence encoding
CDNA of Human F-box
Human CDNA clone (
Novel human coding
Human secreted pro
Human polynucleoti
Human cell growth,
Human CDNA sequenc

9	224.4	80.7	2770	24	ABL99954	F-box protein 44.3
10	213	76.6	1301	22	AAK51857	Human polynucleoti
11	128.2	46.1	406	22	AAF66787	Novel human polynu
12	113.6	40.9	2086	20	AAK85011	Human secreted pro
13	67	24.1	748	20	AAZ15456	Human gene express
14	67	24.1	748	20	AAK98815	Human validated ca
15	67	24.1	754	20	AAZ15699	Human gene express
16	67	24.1	754	20	AAK98893	Human validated ca
17	66	23.7	300	20	AAZ13817	Human gene express
18	66	23.7	300	20	AAK98457	Human cancer cell
19	41.4	14.9	4403765	22	AAI99683	Mycobacterium tube
20	41.4	14.9	4411529	22	AAI99682	Mycobacterium tube
21	39.2	14.1	1312	22	AAO5607	Human secreted pro
22	39.2	14.1	2670	22	AAO5579	Human secreted pro
23	37.6	13.5	698	22	AAZ27215	CDNA encoding nove
24	37.6	13.5	698	22	AAZ27527	CDNA encoding nove
25	37.6	13.5	698	22	AAK55933	Human immune/haema
26	37.6	13.5	1471	22	AAK51721	Human polynucleoti
27	37.6	13.5	1566	22	AAK52704	Human polynucleoti
28	37.6	13.5	1566	22	AAK52705	Human polynucleoti
29	37.6	13.5	1566	24	AAZ24784	Human secreted pro
30	37.6	13.5	1584	21	AAZ56884	Human SBPSAPL poly
31	37.6	13.5	1948	22	AAK51720	Human polynucleoti
32	37.6	13.5	2405	21	AAZ56885	Human SBPSAPL poly
33	37.6	13.5	103599	23	ABX04971	S. cinamonensis m
34	37.4	13.5	135638	25	ABX34289	S. atroolivaceus l
35	36.2	13.0	700	22	AAH92922	Human inflammatory
36	36.2	13.0	1476	24	ABZ74482	Maize peroxidase g
37	36	12.9	1563	22	AAH21494	Human pro-saposin
38	36	12.9	2097	24	ABA94702	Human lipid metabo
39	35.6	12.8	3708	24	AAH47423	Human SLO2 coding
40	35.6	12.8	43280	18	AAH80413	Tylactone synthase
41	35	12.6	1242	25	ABT32145	Benzodiazepines bi
42	35	12.6	32539	25	ABT32129	Benzodiazepines bi
43	34.8	12.5	2241	23	ABL04769	Drosophila melanog
44	34.8	12.5	4864	23	ABL04768	Drosophila melanog
45	34.8	12.5	1403765	22	AAI99683	Mycobacterium tube

ALIGNMENTS

RESULT 1
AAZ93360
ID AAZ93360 standard: DNA; 278 BP.

XX

AC AAZ93360;

XX 16-AUG-2000 (first entry)

DT Sequence encoding F-box protein FBP-10.

XX F-box protein; FBP; diagnosis; treatment; screening; agonist;
XX antagonist; proliferative disorder; differentiation disorder;
XX breast cancer; prostate cancer; ovarian cancer; cancer;
XX small cell lung carcinoma; immune disorder; cardiovascular disorder;
XX inflammatory disorder; human; ss.

OS Homo sapiens.

XX Key Location/Qualifiers
XX CDS 2..277

FT /tag= a

FT /product= FBP-10

FT /note= "This sequence does not encode the
corresponding protein given in AAY83078
unless the stop codon featured below
is removed"

FT misc_feature 44..46

FT /tag= b

FT /note= "TAG stop codon"

XX WO200012679-A1.

xx 09-MAR-2000. 99WO-US19560.
 xx 27-AUG-1999; 99WO-US19560.
 xx 28-AUG-1998; 98US-0098355.
 xx 03-FEB-1999; 99US-0118568.
 xx 15-MAR-1999; 99US-0124449.
 xx (UUNY) UNIV NEW YORK STATE.
 xx Chiau DS, Pagano M, Latres E;
 xx WPI; 2000-256635/22.
 xx P-PSDB; AAY83078.
 xx Novel nucleic acid for screening compounds useful for treating
 PT proliferative and differentiative disorders such as cancer and immune
 PT disorders comprises sequences encoding ubiquitin ligases -
 xx Claim 4; Figure 13b; 245pp; English.
 xx Nucleic acids encoding substrate-targeting subunits of ubiquitin
 CC ligases with F-box motifs (F-box proteins) are useful for diagnosis
 CC of proliferative and differentiated related disorders by measuring
 CC FBP gene expression. Cells expressing such proteins or
 CC their fragments are useful for screening compounds. The compounds
 CC are agonists or antagonists, which are useful for treating a
 CC proliferative or differentiative disorder in a mammal such as
 CC breast, ovarian and prostate cancer and small cell lung carcinoma
 CC and also major opportunistic infections, immune disorders,
 CC cardiovascular diseases and inflammatory disorders. FBP protein,
 CC analogs, derivatives and their subsequences, anti-FBP antibodies
 CC are also useful in diagnosis of the disorders.
 xx Sequence 278 BP; 36 A; 74 C; 107 G; 56 T; 5 other;
 SQ
 Query Match 98.2%; Score 273; DB 21; Length 278;
 Best Local Similarity 100.0%; Pred. No. 5.4e-58;
 Matches 278; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CCGTAGTACTGGNTCCGGGGGCTGGTGGAGGAATGGAGCGGTAGTCTTGGCGCGAG 60
 Db 1 CCGTAGTACTGGNTCCGGGGGCTGGTGGAGGAATGGAGCGGTAGTCTTGGCGCGAG 60
 QY 61 TCCCGGGNTCCCTCCGTAGACCCGCGGANACCTTCGTGTTGAGTAACCTGCGGAGGTGT 120
 Db 61 TCCCGGGNTCCCTCCGTAGACCCGCGGANACCTTCGTGTTGAGTAACCTGCGGAGGTGT 120
 QY 121 GGAGCGTGTGCTACCTTCTCCGCCCAAGCGGTTGCTGCGGGTGGCCTGGCTGTGCGG 180
 Db 121 GGAGCGTGTGCTACCTTCTCCGCCCAAGCGGTTGCTGCGGGTGGCCTGGCTGTGCGG 180
 QY 181 CTTATGGAGGAGTGTGCGGAGAGTATTGGGACCCATCGGAGCGTAACCTGGATCTC 240
 Db 181 CTTATGGAGGAGTGTGCGGAGAGTATTGGGACCCATCGGAGCGTAACCTGGATCTC 240
 QY 241 CGCAGGCGTGGCGGAGCGCGCCACCTGNGGGGCATT 278
 Db 241 CGCAGGCGTGGCGGAGCGCGCCACCTGNGGGGCATT 278
 RESULT 2
 AAL41051
 ID AAL41051 standard; cDNA; 278 BP.
 xx AAL41051;
 AC AAL41051;
 xx 11-OCT-2002 (first entry)
 xx cDNA of Human F-box protein FBP10 SEQ ID No 29.
 xx Cytostatic; immunomodulator; cardiant; antiinflammatory; antimicrobial;

xx proliferative; differentiative disorder; Skp2; F-box protein; cancer;
 KW ubiquitin ligase; breast cancer; prostate cancer; ovarian cancer;
 KW small cell lung carcinoma; immune disorder; parathyroid adenoma; FBP;
 KW inflammatory disorder; lymphoma; major opportunistic infection;
 KW certain cardiovascular disease; human; gene; ss.
 OS Homo sapiens.
 xx WO200255665-A2.
 xx 18-JUL-2002.
 xx 07-JAN-2002; 2002WO-US00311.
 xx 05-JAN-2001; 2001US-260179P.
 xx (UUNY) UNIV NEW YORK STATE.
 xx Pagano M;
 xx WPI; 2002-599665/64.
 xx P-PSDB; AAO22464.
 xx Screening compounds for treating proliferative disorders, e.g. breast
 PT cancer or prostate cancer, infections or immune disorders, comprises
 PT detecting a change in the activity of Skp2 with either p27 or Cks1 -
 xx Disclosure; Fig 13; 246pp; English.
 xx The invention relates to screening compounds useful for the treatment of
 CC proliferative or differentiative disorders comprising detecting a change
 CC in the activity of Skp2 (F-box protein). The method is useful for
 CC screening compounds for the treatment of proliferative or differentiative
 CC disorders, particularly cancer. These compounds include small molecules,
 CC or compounds or derivatives or analogues of the new ubiquitin ligases.
 CC The compounds are useful for treating diseases such as cancer (e.g.
 CC breast cancer, prostate cancer or ovarian cancer, lymphoma, small cell
 CC lung carcinoma or parathyroid adenomas), major opportunistic infections,
 CC immune disorders, certain cardiovascular diseases or inflammatory
 CC disorders. This polynucleotide sequence represents the cDNA encoding an
 CC F-box protein (FBP) of the invention.
 xx Sequence 278 BP; 36 A; 74 C; 107 G; 56 T; 5 other;
 SQ
 Query Match 98.2%; Score 273; DB 24; Length 278;
 Best Local Similarity 100.0%; Pred. No. 5.4e-58;
 Matches 278; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CCGTAGTACTGGNTCCGGCGGGCTGGTGGAGGAATGGAGCGGTAGTCTTGGCGCGAG 60
 Db 1 CCGTAGTACTGGNTCCGGCGGGCTGGTGGAGGAATGGAGCGGTAGTCTTGGCGCGAG 60
 QY 61 TCCCGGGNTCCCTCCGTAGACCCGCGGANACCTTCGTGTTGAGTAACCTGCGGAGGTGT 120
 Db 61 TCCCGGGNTCCCTCCGTAGACCCGCGGANACCTTCGTGTTGAGTAACCTGCGGAGGTGT 120
 QY 121 GGAGCGTGTGCTACCTTCTCCGCCCAAGCGGTTGCTGCGGGTGGCCTGGCTGTGCGG 180
 Db 121 GGAGCGTGTGCTACCTTCTCCGCCCAAGCGGTTGCTGCGGGTGGCCTGGCTGTGCGG 180
 QY 181 CTTATGGAGGAGTGTGTCGCGAGAGTATTGGGACCCATCGGAGCGTAACCTGGATCTC 240
 Db 181 CTTATGGAGGAGTGTGTCGCGAGAGTATTGGGACCCATCGGAGCGTAACCTGGATCTC 240
 QY 241 CGCAGGCGTGGCGGAGCGCGCCACCTGNGGGGCATT 278
 Db 241 CGCAGGCGTGGCGGAGCGCGCCACCTGNGGGGCATT 278
 RESULT 3
 AAH08680
 ID AAH08680 standard; cDNA; 792 BP.
 xx

AAH08680;
 26-JUN-2001 (first entry)
 Human cDNA clone (5'-primer) SEQ ID NO:5515.
 Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
 Homo sapiens.
 EP1074617-A2.
 07-FEB-2001.
 28-JUL-2000; 2000EP-0116126.
 29-JUL-1999; 99JP-0248036.
 27-AUG-1999; 99JP-0300253.
 11-JAN-2000; 2000JP-0118776.
 02-MAY-2000; 2000JP-0183767.
 09-JUN-2000; 2000JP-0241899.
 (HELI-) HELIX RES INST.
 Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 WPI; 2001-318749/34.
 Primer sets for synthesizing polynucleotides, particularly the 5602
 full-length cDNAs defined in the specification, and for the detection
 and/or diagnosis of the abnormality of the proteins encoded by the
 full-length cDNAs -
 Claim 1; SEQ ID 5515; 2537pp + CD ROM; English.
 The present invention describes primer sets for synthesizing 5602
 full-length cDNAs defined in the specification. Where a primer set
 comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 to the complementary strand of a polynucleotide which comprises one of
 the 5602 nucleotide sequences defined in the specification, where the
 oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 of an oligonucleotide comprising a sequence complementary to the
 complementary strand of a polynucleotide which comprises a 5'-end
 sequence and an oligonucleotide comprising a sequence complementary to a
 polynucleotide which comprises a 3'-end sequence, where the
 oligonucleotide comprises at least 15 nucleotides and the combination of
 the 5'-end sequence/3'-end sequence is selected from those defined in
 the specification. The primer sets can be used in antisense therapy and
 in gene therapy. The primers are useful for synthesizing polynucleotides,
 particularly full-length cDNAs. The primers are also useful for the
 detection and/or diagnosis of the abnormality of the proteins encoded by
 the full-length cDNAs. The primers allow obtaining of the full-length
 cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to
 AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632
 represent oligonucleotides, all of which are used in the exemplification
 of the present invention.
 Sequence 792 BP; 183 A; 180 C; 235 G; 191 T; 3 other;
 Query Match 83.5%; Score 232; DB 22; Length 792;
 Best Local Similarity 95.5%; Pred. No. 6.9e-48;
 Matches 257; Conservative 0; Mismatches 10; Indels 2; Gaps 2;

QY 130 GCTCACCTTCCTGCCCGCCCAAGCGGTGCTGCGGGTGCCTGCGCTTATGAG 189
 DB 185 GCTCACCTTCCTGCCCGCCCAAGCGGTGCTGCGGGTGCCTGCGCTTATGAG 244
 QY 190 GGAGTGTGTCGGCAGAGTATTGGGACCCATCGGAGCGTAACCTGGGATCTCCGAGGCCT 249
 DB 245 GGAGTGTGTCGGCAGAGTATTGGGACCCATCGGAGCGTAACCTGGGATCTCCGAGGCCT 304
 QY 250 GCGGAGGTCGGGCCACCTGNGGGGCATT 278
 DB 305 GCGGAGGTCGGGCCACCTGNGGGGCATT 333
 RESULT 4
 ABN59644
 ID ABN59644 standard; cDNA; 1298 BP.
 AC ABN59644;
 XX 28-JUN-2002 (first entry)
 DT Novel human coding sequence SEQ ID NO: 55.
 XX Human; antianaemic; vulnerary; antiinflammatory; immunomodulator;
 KW antifertility; cerebroprotective; cytostatic; rheumatic; gene therapy;
 KW neuroprotective; antiparkinsonian; protein therapy; EST;
 KW expressed sequence tag; gene; ss.
 XX Homo sapiens.
 OS WO200222660-A2.
 PN 21-MAR-2002.
 XX 10-SEP-2001; 2001WO-US26015.
 PF 11-SEP-2000; 2000US-0659671.
 PR (HYSE-) HYSEQ INC.
 PA Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
 PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;
 PI WPI; 2002-292408/33.
 DR P-PSDB; ABB97231.
 XX An isolated polynucleotide for treating diseases associated with its
 PT encoded polypeptide such as cancer and multiple sclerosis -
 XX Claim 1; SEQ ID NO 55; 509pp; English.
 CC The present invention provides the protein and coding sequences of 444
 CC novel human proteins. These were isolated from expressed sequences tags
 CC (ESTs). They can be used to stimulate cell growth, to regulate
 CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
 CC e.g. in burn treatment, to regulate the immune system e.g. to treat
 CC multiple sclerosis, to regulate activin or inhibin e.g. to treat
 CC infertility, to regulate haemostasis or thrombolysis e.g. to treat
 CC stroke and cancer, to screen for drugs, to treat inflammatory conditions
 CC e.g. rheumatoid arthritis, and to treat nervous system disorders e.g.
 CC Parkinson's disease. The present sequence is a coding sequence of the
 CC invention.
 XX Sequence 1298 BP; 337 A; 268 C; 353 G; 340 T; 0 other;
 SQ Query Match 83.5%; Score 232; DB 24; Length 1298;
 Best Local Similarity 95.5%; Pred. No. 7.2e-48;
 Matches 257; Conservative 0; Mismatches 10; Indels 2; Gaps 2;

QY 12 GNTTCCGGGCGGCTGGTGGAGGAATGGAGCCGGTAGTCTTGGCGGAGTCCCG-GGNTC 70
 DB 26 GGTTCGGGCGGCGTGGTGGAGGAATGGAGCCGGTAGTCTTGGCGGAGTCCCGGCGTC 85

QY 71 CTCGCTAGACCCCGGA-NACCTTCGTGTTGAGTAACCTGGCGAGGTGCTGGAGCGTGT 129
 Db 86 CTCGCTAGACCCCGGAGCACCTTCGTGTTGAGTAACCTGGCGAGGTGCTGGAGCGTGT 145
 QY 130 GCTCACCCTTCCTCCGCCAAGCGCTTGTGTCGGGTGGCGTGGCTGTCGGCTTATGGAG 189
 Db 146 GCTCACCCTTCCTCCGCCAAGCGCTTGTGTCGGGTGGCGTGGCTGTCGGCTTATGGAG 205
 QY 190 GGAGTGTGTCGCAGAGTATTGGGACCATCGGACGCTAACCTGATCTCCGAGGCCT 249
 Db 206 GGAGTGTGTCGCAGAGTATTGGGACCATCGGACGCTAACCTGATCTCCGAGGCCT 265
 QY 250 GCGGAGGCGGCCACCTCGGNGGGCATT 278
 Db 266 GCGGAGGCGGCCACCTCGGAGGGCATT 294

RESULT 5

ABA08916

ID ABA08916 standard; cDNA; 1306 BP.

XX ABA08916;

XX ABA08916;

DT 11-JAN-2002 (first entry)

DE Human secreted protein homologue-encoding cDNA, SEQ ID NO:692.

XX Human; cytokine; cell proliferation; cell differentiation; growth factor;
 KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
 KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
 KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
 KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
 KW chronic inflammatory condition; proliferative retinopathy;
 KW atherosclerosis; coronary heart disease; arterial ischaemia;
 KW bone disorder; osteoporosis; vascular growth disorder;
 KW tissue regeneration; wound healing; infection; immune disorder;
 KW cell culture; drug screening; gene therapy; antiinflammatory;
 KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
 KW cytotatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
 KW antifungal; vulnery; antiulcer; ss.

XX Homo sapiens.

XX WO200157188-A2.

XX 09-AUG-2001.

XX 05-FEB-2001; 2001WO-US03800.

XX 03-FEB-2000; 2000US-0496914.

XX 27-APR-2000; 2000US-0560875.

PA (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

XX WPI: 2001-457740/49.

DR P-PSDB; ABB11672.

PT Human proteins and DNA encoding sequences useful for preventing,
 PT treating or ameliorating a medical condition in a mammalian subject
 PT e.g. arthritis and cancer -

PS Claim 1; Page 644; 1963pp; English.

CC Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
 CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
 CC invention also relates to vectors and recombinant host cells comprising a
 CC nucleotide of the invention, methods of producing the novel polypeptides,
 CC antibodies against the polypeptides, methods of detecting the nucleotides
 CC or polypeptides in a sample, and methods of identifying compounds which
 CC bind to polypeptides of the invention. Although novel, many of the
 CC polypeptides of the invention have homology to known proteins, thereby

CC giving an insight into their probable biological activities, and hence
 CC potential therapeutic applications. The polypeptides of the invention may
 CC have various activities, including cytokine, cell proliferation or cell
 CC differentiation activities; stem cell growth factor activity;
 CC haematopoiesis regulatory activity; tissue growth activity;
 CC immunomodulatory activity; activin- or inhibin-related activities;
 CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
 CC thrombolytic activities; receptor or ligand activities; or may be
 CC involved in oncogenesis, cancer cell proliferation or metastasis.
 CC Depending on their biological activities, polypeptides and nucleotides of
 CC the invention are useful for preventing, treating or ameliorating medical
 CC conditions, e.g., by protein or gene therapy. Such conditions include
 CC cancers, haematopoietic disorders (e.g. myeloid or lymphoid cell
 CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
 CC proliferative retinopathy, atherosclerosis, coronary heart disease,
 CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
 CC vascular growth. Polypeptides involved with tissue regeneration and
 CC repair (or nucleic acids encoding them) may be used to promote wound
 CC healing (e.g., of burns, incisions and ulcers), while those with
 CC immunomodulatory activities may be used in the treatment of viral,
 CC bacterial and fungal infections in addition to immune disorders.
 CC Polypeptides with growth factor activity may be used in cell cultures to
 CC promote cell growth. For example, such polypeptides may be used to
 CC manipulate stem cells in culture to give rise to neuroepithelial cells
 CC that can be used to augment or replace cells damaged by illness,
 CC autoimmune disease or accidental damage. The polypeptides and nucleotides
 CC may also be used in the diagnosis of the above conditions, and in drug
 CC screening techniques. The present sequence represents a cDNA encoding a
 CC novel human polypeptide of the invention.

XX Sequence 1306 BP; 339 A; 269 C; 355 G; 343 T; 0 other;

Query Match

Best Local Similarity 83.5%; Score 232; DB 22; Length 1306;

Matches 257; Conservative 0; Mismatches 10; Indels 2; Gaps 2;

QY 12 GNTTCCGCGCGCTGTGAGGAATGGAGCCGGTAGTCTTCGCGGAGTCCCG-GGNTC 70
 Db 37 GTTTCGCGCGCGCTGTGAGGAATGGAGCCGGTAGTCTTCGCGGAGTCCCGCGGCTC 96
 QY 71 CTCGCTAGACCCCGGA-NACCTTCGTGTTGAGTAACCTGGCGAGGTGCTGGAGCGTGT 129
 Db 97 CTCGCTAGACCCCGGAGCACCTTCGTGTTGAGTAACCTGGCGAGGTGCTGGAGCGTGT 156
 QY 130 GCTCACCCTTCCTCCGCCAAGCGCTTGTGTCGGGTGGCGCTGCTGCTTATGGAG 189
 Db 157 GCTCACCCTTCCTCCGCCAAGCGCTTGTGTCGGGTGGCGCTGCTGCTTATGGAG 216
 QY 190 GGAGTGTGTCGCAGAGTATTGGGACCATCGGACGCTAACCTGATCTCCGAGGCCT 249
 Db 217 GGAGTGTGTCGCAGAGTATTGGGACCATCGGACGCTAACCTGATCTCCGAGGCCT 276
 QY 250 GCGGAGGCGGCCACCTCGGNGGGCATT 278
 Db 277 GCGGAGGCGGCCACCTCGGAGGGCATT 305

RESULT 6

AAK52841

ID AAK52841 standard; cDNA; 1306 BP.

XX AAK52841;

XX AAK52841;

XX 06-NOV-2001 (first entry)

DT Human polynucleotide SEQ ID NO 2370.

DE Human; cytokine; cell proliferation; cell differentiation; gene therapy;

XX vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation; ss.
 XX Homo sapiens.

XX WO200157190-A2.
 XX 09-AUG-2001.
 XX 05-FEB-2001; 2001WO-US04098.
 XX 03-FEB-2000; 2000US-0496914.
 XX 27-APR-2000; 2000US-0560875.
 XX 20-JUN-2000; 2000US-0598075.
 XX 19-JUL-2000; 2000US-0620325.
 XX 01-SEP-2000; 2000US-0654936.
 XX 15-SEP-2000; 2000US-0663561.
 XX 20-OCT-2000; 2000US-0693325.
 XX 30-NOV-2000; 2000US-0728422.
 XX (HYSE-) HYSEQ INC.
 XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
 XX Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 XX Xue AJ, Yang Y, Wejhrman T, Goodrich R;
 XX WPI: 2001-476283/51.
 XX P-PSDB: AM79708.
 XX Nucleic acids encoding polypeptides with cytokine-like activities,
 XX useful in diagnosis and gene therapy -
 XX Claim 1; Page 4663-4664; 6221pp; English.
 XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
 XX encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
 XX cytokine, cell proliferation or cell differentiation or which may induce
 XX production of other cytokines in other cell populations. The
 XX polynucleotides and polypeptides are useful in gene therapy, vaccines or
 XX peptide therapy. The polypeptides have various cytokine-like activities,
 XX e.g. stem cell growth factor activity, haematopoiesis regulating
 XX activity, tissue growth factor activity, immunomodulatory activity and
 XX activin/inhibin activity and may be useful in the diagnosis and/or
 XX treatment of cancer, leukaemia, nervous system disorders, arthritis and
 XX inflammation.
 XX Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
 XX (AAM80020) are omitted as the relevant pages from the sequence listing
 XX were missing at the time of publication.
 XX Sequence 1306 BP; 339 A; 269 C; 355 G; 343 T; 0 other;
 XX
 XX Query Match 83.5%; Score 232; DB 22; Length 1306;
 XX Best Local Similarity 95.5%; Pred. No. 7.2e-48;
 XX Matches 257; Conservative 0; Mismatches 10; Indels 2; Gaps 2;
 QY 12 GNTTCCGGCGGCTGTGAGGAATGAGCGGTAGTGTCTGCGCGAGTCCCG-GGNTC 70
 Db 37 GGTTCGGCGGCTGTGAGGAATGAGCGGTAGTGTCTGCGCGAGTCCCGCGGCTC 96
 QY 71 CTCCTGAGACCCCGGA-NACCTTCGTGTGAGTAACCTGGCGAGGTGTGGAGCGTGT 129
 Db 97 CTCCTGAGACCCCGGAGCACCTTCGTGTGAGTAACCTGGCGAGGTGTGGAGCGTGT 156
 QY 130 GCTCACCCTTCCTGCCCGCAAGCGTGTCTGCGGGTGGCGTGGCGTGTATGGAG 189
 Db 157 GTCACCTTCCTGCCCGCAAGCGTGTCTGCGGGTGGCGTGGCGTGTATGGAG 216
 QY 190 GGAGTGTGTGCGCAGAGTATTGCGGACCATCGGAGCGTAACCTGGATCTCCGAGCGCT 249
 Db 217 GGAGTGTGTGCGCAGAGTATTGCGGACCATCGGAGCGTAACCTGGATCTCCGAGCGCT 276
 QY 250 GCGGAGGCGGCGCACCTTCGNGGGGCATT 278
 Db 277 GCGGAGGCGGCGCACCTTCGAGGGGCATT 305

RESULT 7

ABZ24703
 ID ABZ24703 standard; cDNA; 1418 BP.
 XX AC ABZ24703;
 XX DT 07-APR-2003 (first entry)
 XX Human cell growth, differentiation and death protein CGDD-15 cDNA.
 DE CGDD-15; cell growth; cell differentiation; hepatotropic; cell death; human;
 KW cytosolic; antiarteriosclerotic; hepatotropic; antiinflammatory;
 KW antiproliferative; antianemic; ophthalmological; auditory;
 KW anticonvulsant; cerebroprotective; nootropic; neuroprotective;
 KW antiparkinsonian; neuroleptic; tranquilizer; immunosuppressive;
 KW anti-HIV; anti-allergic; antilasthmatic; antithyroid; antidiabetic;
 KW dermatological; nephrotropic; antirheumatic; antiarthritic;
 KW antitumor; vulnery; virucide; antibacterial; fungicide;
 KW antiparasitic; protozoacide; antihelminthic; antifertility;
 KW gynaecological; gene therapy; microarray; gene; ss.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 PH 99..1310
 FT /*tag= a
 FT /product= "Human CGDD-15"
 XX WO200297032-A2.
 XX 05-DEC-2002.
 XX 05-APR-2002; 2002WO-US11152.
 XX 06-APR-2001; 2001US-282110P.
 XX 11-APR-2001; 2001US-283294P.
 XX 26-APR-2001; 2001US-286820P.
 XX 27-APR-2001; 2001US-287228P.
 XX 16-MAY-2001; 2001US-291662P.
 XX 18-MAY-2001; 2001US-291846P.
 XX 25-MAY-2001; 2001US-293727P.
 XX 01-JUN-2001; 2001US-295263P.
 XX 01-JUN-2001; 2001US-295340P.
 XX 15-JAN-2002; 2002US-349705P.
 XX (INCY-) INCYTE GENOMICS INC.
 XX Azimzai Y, Au-Young JK, Batra S, Baughn MR, Becha SD, Borowsky ML;
 XX Burford N, Ding L, Elliott VS, Emerling BM, Gandhi AR, Gietzen KJ;
 XX Griffin JA, Hafalia AJA, Honchell CD, Lal PG, Lee SY, Lu DAM;
 XX Arvizu CS, Rankumar J, Reddy R, Sanjanwala MM, Tang YT, Walla NK;
 XX Wang YE, Warren BA, Xu Y, Yang J, Yao MG, Yue H, Zebarjadian Y;
 XX WPI: 2003-140453/13.
 XX P-PSDB: ABP58344.
 XX Novel human proteins associated with cell growth, differentiation and
 XX death, useful for treating, diagnosing or preventing cancer,
 XX developmental, neurological, reproductive or autoimmune/inflammatory
 XX disorders
 XX Claim 5; Page 230-231; 238pp; English.
 XX The present sequence is that of incyte clone 5565648CB1 encoding
 XX human CGDD-15, a novel protein associated with cell growth,
 XX differentiation and death. A representative cDNA library for the
 XX polynucleotide is LIVERF05 from foetal liver tissue RNA.
 XX Structural features establish the encoded protein as being
 XX associated with cell growth, differentiation and death. The
 XX invention is based on novel human CGDD-1 to -21 proteins (see
 XX ABP58330-50), the polynucleotides encoding them (see ABZ24689-709),
 XX and to the use of these for the diagnosis, treatment or prevention
 XX of cell proliferative disorders including cancer, developmental
 XX disorders, neurological disorders, autoimmune disorders,

CC reproductive disorders, and disorders of the placenta, and in the
 CC assessment of the effects of exogenous compounds on the activity
 CC and expression of proteins and nucleic acids associated with cell
 CC growth, differentiation and death. CGDD polynucleotides are also
 CC used in a claimed microarray and in a claimed method of generating
 CC an expression profile of a sample.

XX
 SQ Sequence 1418 BP; 355 A; 298 C; 384 G; 381 T; 0 other;

Query Match 83.5%; Score 232; DB 25; Length 1418;
 Best Local Similarity 95.5%; Pred. No. 7.3e-48;
 Matches 257; Conservative 0; Mismatches 10; Indels 2; Gaps 2;
 QY 12 GNTTCGCGCGGCTGGTGGAGGATGAGCGCGGTAGTGTGGCGGAGTCCCG-GGNTC 70
 Db. 77 GGTTCGCGCGGCTGGTGGAGGATGAGCGCGGTAGTGTGGCGGAGTCCCGGCGCTC 136
 QY 71 CTCGCTAGACCCCGGGA-NACCTTCGTGTGTAGTAACCTGGCGGAGGTGGTGGAGCGTGT 129
 Db 137 CTCGCTAGACCCCGGAGCACCTTCGTGTGTAGTAACCTGGCGGAGGTGGTGGAGCGTGT 196
 QY 130 GCTCACCCTTCCTCCCGCCCAAGCGTGTGCTGGCGGTGGCTGCGTGTCCGCTTATGGAG 189
 Db 197 GCTCACCCTTCCTCCCGCCCAAGCGTGTGCTGGCGGTGGCTGCGTGTCCGCTTATGGAG 256
 QY. 190 GGAGTGTGTGGCAGAGTATTGGCGACCCATCGGACGCGTAACCTGGATCTCCGCGAGCCT 249
 Db 257 GGAGTGTGTGGCAGAGTATTGGCGACCCATCGGACGCGTAACCTGGATCTCCGCGAGCCT 316
 QY 250 GCGCGAGGCGCGCCACCTCGNGGGGCATT 278
 Db 317 GCGCGAGGCGCGCCACCTCGAGGGGCATT 345

RESULT 8
 AAHL16815 standard; cDNA; 1486 BP.
 XX
 AC AAHL16815;
 XX
 DT 26-JUN-2001 (first entry)
 XX
 DE Human cDNA sequence SEQ ID NO:16073.
 XX
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
 XX
 OS Homo sapiens.
 XX
 PN EP1074617-A2.
 XX
 PD 07-FEB-2001.
 XX
 PF 28-JUL-2000; 2000EP-0116126.
 XX
 PR 29-JUL-1999; 99JP-0248036.
 PR 27-AUG-1999; 99JP-0300253.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 DR WPI; 2001-318749/34.
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs.
 XX
 PS Claim 8; SEQ ID 16073; 2537pp + CD ROM; English.

XX
 CC The present invention describes primer sets for synthesising 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesising polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.

XX Sequence 1486 BP; 369 A; 301 C; 367 G; 449 T; 0 other;

Query Match 83.5%; Score 232; DB 22; Length 1486;
 Best Local Similarity 95.5%; Pred. No. 7.3e-48;
 Matches 257; Conservative 0; Mismatches 10; Indels 2; Gaps 2;
 QY 12 GNTTCGCGCGGCTGGTGGAGGATGAGCGCGGTAGTGTGGCGGAGTCCCG-GGNTC 70
 Db 65 GGTTCGCGCGGCTGGTGGAGGATGAGCGCGGTAGTGTGGCGGAGTCCCGGCGCTC 124
 QY 71 CTCGCTAGACCCCGGGA-NACCTTCGTGTGTAGTAACCTGGCGGAGGTGGTGGAGCGTGT 129
 Db 125 CTCGCTAGACCCCGGAGCACCTTCGTGTGTAGTAACCTGGCGGAGGTGGTGGAGCGTGT 184
 QY 130 GCTCACCCTTCCTCCCGCCCAAGCGTGTGCTGGCGGTGGCTGCGTGTCCGCTTATGGAG 189
 Db 185 GCTCACCCTTCCTCCCGCCCAAGCGTGTGCTGGCGGTGGCTGCGTGTCCGCTTATGGAG 244
 QY 190 GGAGTGTGTGGCAGAGTATTGGCGACCCATCGGACGCGTAACCTGGATCTCCGCGAGCCT 249
 Db 245 GGAGTGTGTGGCAGAGTATTGGCGACCCATCGGACGCGTAACCTGGATCTCCGCGAGCCT 304
 QY 250 GCGCGAGGCGCGCCACCTCGNGGGGCATT 278
 Db 305 GCGCGAGGCGCGCCACCTCGAGGGGCATT 333

RESULT 9
 ABL99954 standard; cDNA; 2770 BP.
 ID ABL99954
 XX
 AC ABL99954;
 XX
 DT 08-AUG-2002 (first entry)
 XX
 DE F-box protein 44.33 encoding cDNA SEQ ID NO 1.
 XX
 KW F-box protein; cytostatic; virucidal; immunomodulatory; HIV;
 KW antinflammatory; haemostatic; malignant tumour; infection;
 KW human immunodeficiency virus; immunological disease; gene therapy;
 KW gene; ss.
 XX
 OS Unidentified.
 XX
 FH Key Location/Qualifiers
 FT CDS 16..1227
 FT /tag= a
 FT /product= "F-box protein 44.33"
 FT /note= "claimed in claim 6"

XX PN WO200232951-A1.
 XX PD 25-APR-2002.
 XX PF 21-SEP-2001; 2001WO-CN01443.
 XX PR 22-SEP-2000; 2000CN-0125325.
 XX PA (BIOW-) BIOWINDOW GENE DEV INC SHANGHAI.
 XX PI Mao Y, Xie Y;
 XX DR WPI; 2002-340233/37.
 XX P-PSDB; ABB77551.
 XX PT A novel polypeptide, a F-box protein 44.33 and encoding polynucleotide,
 XX used in diagnosis and treatment of diabetes, menstrual disturbance,
 XX peptic ulcer, arrhythmia, hemophthisis and epilepsy -
 XX PS Claim 6; Page 26-27; 36pp; Chinese.
 XX CC The invention relates to F-box protein 44.33 with cytostatic,
 XX virucidal, immunomodulatory, antiinflammatory and haemostatic
 XX activity. The protein and encoding polynucleotide are used in diagnosis
 XX and treatment of malignant tumour, haemopathy, human immunodeficiency
 XX virus (HIV) infection, immunological diseases and various inflammations.
 XX CC The polynucleotide is useful in gene therapy.
 XX SQ Sequence 2770 BP; 794 A; 509 C; 613 G; 854 T; 0 other;
 Query Match 80.7%; Score 224.4; DB 24; Length 2770;
 Best Local Similarity 95.4%; Pred. No. 5.7e-46;
 Matches 250; Conservative 0; Mismatches 10; Indels 2; Gaps 2;
 Qy 19 GCGGGCTGGTGAAGAATGGAGCCGGTAGTGTGCGGCGAGTCCCG-GGNTCTCCGTA 77
 Db 1 GGGGGCTGGTGAAGAATGGAGCCGGTAGTGTGCGGCGAGTCCCGGCTCCTCCGTA 60
 Qy 78 GACCCGCGGA-NACCTTCGTGTTGAGTAACCTGCGGAGGTGGTGGAGCGTGTCTCACC 136
 Db 61 GACCCGCGGACCTTCGTGTTGAGTAACCTGCGGAGGTGGTGGAGCGTGTCTCACC 120
 Qy 137 TTCCTGCCCGCCAAAGGCGTGTGCTGCGGGTGGCCCTGCTGTGCCCTTATGGAGGAGTGT 196
 Db 121 TTCCTGCCCGCCAAAGGCGTGTGCTGCGGGTGGCCCTGCTGTGCCCTTATGGAGGAGTGT 180
 Qy 197 GTGCGCAGAGTATTGCGGACCCATCGGAGCGGTAACTGGATCTCCGAGGCGCTGGCGGAG 256
 Db 181 GTGCGCAGAGTATTGCGGACCCATCGGAGCGGTAACTGGATCTCCGAGGCGCTGGCGGAG 240
 Qy 257 GCCGGCCACCTGGNGGGGCATT 278
 Db 241 GCCGGCCACCTGGAGGGGCATT 262
 RESULT 10
 AAK51857
 ID AAK51857 standard; cDNA; 1301 BP.
 XX AC AAK51857;
 XX DT 06-NOV-2001 (first entry)
 XX DE Human polynucleotide SEQ ID NO 402.
 XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation; ss.
 OS Homo sapiens.
 XX

PN WO200157190-A2.
 PD 09-AUG-2001.
 PF 05-FEB-2001; 2001WO-US04098.
 PR 03-FEB-2000; 2000US-0496914.
 PR 27-APR-2000; 2000US-0560875.
 PR 20-JUN-2000; 2000US-0598075.
 PR 19-JUL-2000; 2000US-0620325.
 PR 01-SEP-2000; 2000US-0634936.
 PR 15-SEP-2000; 2000US-0663561.
 PR 20-OCT-2000; 2000US-0693325.
 PR 30-NOV-2000; 2000US-0728422.
 XX (HYSE-) HYSEQ INC.
 XX PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
 PI Zhao QA, Wang L, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 PI xue AJ, Yang Y, Wejhrman T, Goodrich R;
 XX DR WPI; 2001-476283/51.
 XX P-PSDB; AAM78724.
 XX PT Nucleic acids encoding polypeptides with cytokine-like activities,
 XX useful in diagnosis and gene therapy -
 XX PS Claim 1; Page 1522-1524; 6221pp; English.
 XX CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
 CC (AAM80020) are omitted as the relevant pages from the sequence listing
 CC were missing at the time of publication.
 XX SQ Sequence 1301 BP, 340 A; 270 C; 350 G; 341 T; 0 other;
 Query Match 76.6%; Score 213; DB 22; Length 1301;
 Best Local Similarity 95.6%; Pred. No. 3.3e-43;
 Matches 238; Conservative 0; Mismatches 9; Indels 2; Gaps 2;
 Qy 32 GAATGGAGCCGCTAGTGTGCGGCGAGTCCCG-GGNTCTCTCCGTAGACCCGCGGA-NA 89
 Db 54 GAATGGAGCCGCTAGTGTGCGGCGAGTCCCGGCGCTCTCCGTAGACCCGCGGACA 113
 Qy 90 CCTTCGTGTTGAGTAACCTGCGGAGGTGGTGGAGCGTGTCTCACCTTCTGCGCCGCA 149
 Db 114 CCTTCGTGTTGAGTAACCTGCGGAGGTGGTGGAGCGTGTCTCACCTTCTGCGCCGCA 173
 Qy 150 AGCGGTTCTCGGGTGGCCCTGCGTTCGCCCTTATGGAGGAGTGTGTGGCAGAGTAT 209
 Db 174 AGCGGTTCTCGGGTGGCCCTGCGTTCGCCCTTATGGAGGAGTGTGTGGCAGAGTAT 233
 Qy 210 TCGGAGCCCATCGGAGCGTAACCTGGATCTCCGAGGCGCTGGCGAGGCGCGCCACCTGG 269
 Db 234 TCGGAGCCCATCGGAGCGTAACCTGGATCTCCGAGGCGCTGGCGAGGCGCGCCACCTGG 293
 Qy 270 NGGGGCATT 278
 Db 294 AGGGGCATT 302
 RESULT 11
 AAF66787

ID AAF66787 standard; cDNA; 406 BP.
 AC AAF66787;
 XX
 DT 09-APR-2001 (first entry)
 XX
 DE Novel human polynucleotide, SEQ ID NO: 2543.
 XX
 KW Human; cytostatic; gene therapy; colon cancer; prostate cancer;
 KW breast cancer; lung cancer; cancer detection; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200102568-A2.
 XX
 PD 11-JAN-2001.
 XX
 PF 30-JUN-2000; 2000WO-US18374.
 XX
 PR 02-JUL-1999; 99US-0142310.
 PR 02-JUL-1999; 99US-0142311.
 XX
 PA (CHIR) CHIRON CORP.
 PA (HYSE-) HYSEQ INC.
 XX
 PI Williams LT, Escobedo J, Innis MA, Garcia PD, Klinger J, Kassam A;
 PI Reinhard C, Randazzo F, Kennedy GC, Pot D, Lamson G, Drmanac R;
 PI Crkenjakov R, Drmanac S, Dickson M, Labat I, Leshkowitz D;
 PI Kita D, Garcia V, Jones LW, Strache-Crain B;
 XX
 DR WPI; 2001-091805/10.
 XX
 PT Library of polynucleotides for diagnosing a cancerous state of a
 PT mammalian cell and detecting cancer, particularly of the colon or
 PT prostate, comprises 3351 human polynucleotide sequences -
 XX
 PS Claim 9; Page 919-920; 1046pp; English.
 XX
 CC The present sequence is one of 3351 sequences in a library of human
 CC polynucleotides. The library is used to detect differentially expressed
 CC genes correlated with a cancerous state of a mammalian cell and can
 CC detect colon, prostate, breast and lung cancer. The library can be used
 CC to produce probes for detection of mRNA and to produce additional copies
 CC of the polynucleotides. The probes can be used for chromosome mapping of
 CC the polynucleotide and for detection of transcription levels. Ribozymes
 CC or antisense oligonucleotides can be generated. The polynucleotides and
 CC their gene products are used as genetic or biochemical markers (e.g. in
 CC blood or tissues) that will detect the earliest changes along the
 CC carcinogenesis pathway and/or monitor the efficacy of therapies and
 CC preventive interventions. The polynucleotides, polypeptides and
 CC antibodies against them can be used in pharmaceutical compositions to
 CC treat the cancers and proliferative disorders such as neoplasia,
 CC dysplasia and hyperplasia.
 XX
 SQ Sequence 406 BP; 90 A; 100 C; 124 G; 90 T; 2 other;

Query Match 46.1%; Score 128.2; DB 22; Length 406;
 Best Local Similarity 91.7%; Pred. No. 2e-22;
 Matches 155; Conservative 0; Mismatches 12; Indels 2; Gaps 2;
 QY 12 GNTTCCGGCGGCTGTGAGGATGAGCGGTAGTGTGTCGGCGGAGTCCCG-GGNTC 70
 Db 29 GGTTCGGCGGCGGCTGTGAGGATGAGCGGTAGTGTGTCGGCGGAGTCCCGGCGTC 88
 QY 71 CTCGCTAGACCCCGGGA-NACCTTCGTGTTGAGTAACCTGCGGAGTGTGAGCGGTG 129
 Db 89 CTCGCTAGACCCCGGAGACCTTCGTGTTGAGTAACCTGCGGAGTGTGAGCGGTG 148
 QY 130 GCTCACCTTCTGCGCCGCCAAGCGGTGTGTCGGGTGGCGTGGCTGTC 178
 Db 149 GCTCACCTTCTGCGCCGCCAAGCGGTGTGTCGGGTGGCGTGGCTGTC 197

RESULT 12

AAX85011
 ID AAX85011 standard; DNA; 2086 BP.
 XX
 AC AAX85011;
 XX
 DT 30-JUL-1999 (first entry)
 XX
 DE Human secreted protein gene No. 79.
 XX
 KW Human; secreted protein; fusion protein; gene therapy; protein therapy;
 KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
 KW developmental abnormality; foetal deficiency; blood; allergy; renal; ds;
 KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
 KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
 KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
 KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
 KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
 XX
 OS Homo sapiens.
 XX
 PN WO9924836-A1.
 XX
 PD 20-MAY-1999.
 XX
 PF 04-NOV-1998; 98WO-US23435.
 XX
 PR 17-NOV-1997; 97US-0066100.
 PR 07-NOV-1997; 97US-0064900.
 PR 07-NOV-1997; 97US-0064908.
 PR 07-NOV-1997; 97US-0064911.
 PR 07-NOV-1997; 97US-0064912.
 PR 07-NOV-1997; 97US-0064983.
 PR 07-NOV-1997; 97US-0064984.
 PR 07-NOV-1997; 97US-0064985.
 PR 07-NOV-1997; 97US-0064987.
 PR 07-NOV-1997; 97US-0064988.
 PR 17-NOV-1997; 97US-0066090.
 PR 17-NOV-1997; 97US-0066094.
 PR 17-NOV-1997; 97US-0066095.
 PR 17-NOV-1997; 97US-0066089.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 CC Carter KC, Ebner R, Endress GA, Feng P, Janat F;
 CC Kyaw H, Lafleur DM, Moore PA, Ni J, Olsen HS, Rosen CA;
 CC Ruben SM, Shi Y, Soppet DR, Wei Y;
 XX
 DR WPI; 1999-337740/28.
 DR P-PSDB: AAY27645, AAY27795, AAY27796, AAY27797, AAY27798, AAY27799,
 DR AAY27800, AAY27801, AAY27802, AAY27803.
 XX
 PT New human secreted proteins and coding sequences useful for treating
 PT disorders of the immune system and hyperproliferative disorders
 XX
 PS Claim 1; Page 323; 507pp; English.
 XX
 CC This sequence represents a nucleic acid molecule which encodes a
 CC secreted human protein. The gene number is given in the descriptor line.
 CC The gene can be used to generate fusion proteins by linking to the gene
 CC to a human immunoglobulin Fc portion (e.g. AAX84924) for increasing the
 CC stability of the fused protein as compared to the human protein only.
 CC The invention relates to 125 novel genes and their fragments (nucleic
 CC acid sequences: AAX84933-X85057; amino acid sequences AAY27567-Y27933)
 CC which are useful for preventing, treating or ameliorating medical
 CC conditions e.g. by protein or gene therapy. Also, pathological
 CC conditions can be diagnosed by determining the amount of the new
 CC polypeptides in a sample or by determining the presence of mutations in
 CC the new polynucleotides. Specific uses are described for each of the 125
 CC polynucleotides, based on which tissues they are most highly expressed in
 CC (see AAX84933 for described uses).
 XX
 SQ Sequence 2086 BP; 611 A; 373 C; 442 G; 656 T; 4 other;

can be used for raising antibodies for experimental, diagnostic and therapeutic purposes. The polynucleotides may also be used to construct arrays for diagnostics (which may be used to determine function of an encoded protein); and to detect differences in expression levels between two cells (e.g. to identify abnormal or diseased tissue in a human, to identify a genetic predisposition or susceptibility to a disease such as cancer). The polynucleotides of the invention are especially used in the diagnosis, prognosis and management of colorectal cancer, breast cancer, and lung cancer. The polynucleotides can also be used to screen for peptide analogues and antagonists.

Sequence 748 BP; 198 A; 152 C; 171 G; 211 T; 16 other;

Query Match 24.1%; Score 67; DB 20; Length 748;
Best Local Similarity 98.5%; Pred. NO. 2.2e-07;
Matches 67; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 211 GCGGACCCATCGGAGCGTAACCTGGATCTCCGAGGCTGCGGAGCGCCGACCTGGN 270
DB 71 GCGGACCCATCGGAGCGTAACCTGGATCTCCGAGGCTGCGGAGCGCCGACCTGGA 130
QY 271 GGGGCATY 278
DB 131 GGGGCATY 138

RESULT 14
AA98815
ID AA98815 standard; cDNA; 748 BP.
XX AC AA98815;
XX DT 24-SEP-1999 (first entry)
XX DE Human validated cancer cell derived cDNA #137.
XX KW Cancer; human; colon; breast; lung; transmembrane receptor; ATPase;
KW integral membrane protein; aspartyl protease; GATA family; wnt family;
KW transcription factor; G-protein alpha subunit; protein phosphatase;
KW phospholipase binding protein; diacylglycerol binding protein; trypsin;
KW protein kinase; tyrosine phosphatase; developmental signalling protein;
KW WW/WWP domain; therapy; forensic; genetic mapping; diagnostic;
KW detection; treatment; cervical; melanoma; colorectal adenocarcinoma;
KW Wilms' tumour; retinoblastoma; sarcoma; myosarcoma; lung carcinoma;
KW leukemia; lymphoma; dysplasia; hyperplasia; endometrium; adrenal;
XX KW prostate; ss.
XX OS Homo sapiens.
XX PN WO9933982-A2.
XX PD 08-JUL-1999.
XX PF 22-DEC-1998; 98WO-US27610.
XX PR 21-DEC-1998; 98JS-0217471.
PR 23-DEC-1997; 97US-0068755.
PR 03-APR-1998; 98JS-0080664.
PR 21-OCT-1998; 98JS-0105234.
PR 27-OCT-1998; 98JS-0105877.
XX (CHIR) CHIRON CORP.
PA (HYSE-) HYSEQ INC.
XX PI Crkvenjakov R, Dickson M, Drmanac R, Drmanac S;
PI Escobedo J, Garcia PD, Garcia V, Giese K, Innis MA;
PI Jones WL, Kassam A, Kennedy GC, Kita D, Labat I;
PI Lamson G, Leshkowitz D, Pot D, Randazzo F, Reinhard C;
PI Stache-Crain B, Sudduth-Klinger J, Williams LT;
XX WPI; 1999-430243/16.
XX New isolated human polynucleotides

Query Match 40.9%; Score 113.6; DB 20; Length 2086;
Best Local Similarity 88.1%; Pred. NO. 8.9e-19;
Matches 133; Conservative 0; Mismatches 17; Indels 1; Gaps 1;

QY 29 GAGGAATGAGCGGTAGTGTCTTGGCGGAGTCCCGGNTCCCTAGACCGCGGA- 87
DB 2 GAGGAATGAGCGGTAGTGTCTTGGCGGAGTCCCGGNTCCCTAGACCGCGGAG 61
QY 88 NACCTTCGTTGTAGTAACCTGGCGGAGTGGTGGAGCGTGTGCTACCTTCTTGCCTGC 147
DB 62 CACCTTCGTTGTAGTAACCTGGCGGAGTGGTGGAGCGTGTGCTACCTTCTTGCCTGC 121
QY 148 CAGGCGTGTGCGGGTGGCGTGGCTGGCTGTGC 178
DB 122 CAGGCGTGTGCGGGTGGCGTGGCTGAATGTTTC 152

RESULT 13
AAZ15456.
ID AAZ15456 standard; cDNA; 748 BP.
XX AC AAZ15456;
XX DT 12-OCT-1999 (first entry)
XX DE Human gene expression product cDNA sequence SEQ ID NO:2925.
XX KW Human; gene; gene expression product; diagnosis; therapy; probe;
KW detection; mapping; tissue typing; profiling; forensic; cancer;
KW genetic analysis; colorectal cancer; breast cancer; lung cancer; ss.
XX OS Homo sapiens.
XX PN WO9938972-A2.
XX PD 05-AUG-1999.
XX PF 28-JAN-1999; 99WO-US01619.
XX PR 03-APR-1998; 98US-0080666.
PR 28-JAN-1998; 98US-0072910.
PR 24-FEB-1998; 98US-0079954.
PR 31-MAR-1998; 98US-0080114.
PR 03-APR-1998; 98US-0080515.
XX (CHIR) CHIRON CORP.
PA (HYSE-) HYSEQ INC.
XX PI Crkvenjakov R, Dickson M, Drmanac R, Drmanac S;
PI Escobedo J, Garcia PD, Garcia V, Giese K, Innis MA;
PI Jones WL, Kassam A, Kennedy GC, Kita D, Labat I;
PI Lamson G, Leshkowitz D, Pot D, Randazzo F, Reinhard C;
PI Stache-Crain B, Sudduth-Klinger J, Williams LT;
XX WPI; 1999-494092/41.
XX Novel human genes and their expression products which are
XX differentially expressed in different cell types
PS Claim 1; Page 1411-1412; 2479pp; English.
XX The present invention describes a library of human polynucleotides
XX comprising the sequences given in AAZ12532 to AAZ17779. Also described is
XX a method of detecting differentially expressed genes correlated with the
XX cancerous state of a mammalian cell, comprising detecting at least one
XX differentially expressed gene product in a test sample from a cell
XX suspected of being cancerous, where the gene product is encoded by one
XX of the 5248 polynucleotide sequences given in AAZ12532 to AAZ17779. The
XX polynucleotides can be used as a source of primers and probes, which can
XX be used for a variety of purpose, e.g. detection of expression levels,
XX mapping, tissue typing or profiling, forensic, genetic analysis and
XX detection of polymorphisms. Polypeptides encoded by the polynucleotides

XX Claim 1; Page 468-469; 591pp; English.

XX This invention describes novel isolated human polynucleotides obtained

CC by screening for differential expression in colon cancer, breast cancer

CC and lung cancer cell lines. The polynucleotides of the invention are

CC represented in AAX98275-X99118 and encode polypeptides of protein

CC families selected from 4 transmembrane segments integral membrane

CC proteins, 7 transmembrane receptors, ATPases associated with various

CC cellular activities (AAA), eukaryotic aspartyl proteases, GATA family of

CC transcription factors, G-protein alpha subunit, phospholipase or

CC diacylglycerol binding proteins, protein kinase, protein phosphatase 2C,

CC protein tyrosine phosphatase, trypsin, wnt family of developmental

CC signalling proteins and WW/SP5/WWP domain containing proteins. The

CC encoded polypeptides also have a functional domain selected from Ank

CC repeat, basic region plus leucine zipper transcription factors,

CC bromodomain, EF-hand, SH3 domain, WD domain/G-beta repeats, zinc finger

CC (C2H2 type), zinc finger (CCHC class), and zinc-binding metalloprotease

CC domain. The polynucleotides encode polypeptides with similarity to known

CC protein families and are predicted to have similar properties. The novel

CC polynucleotides can be used to develop products for use as therapeutic

CC agents and in forensics, genetic analysis, mapping and diagnostic

CC applications. In particular, the product can be used for the detection

CC and management of cancers. They can be used for treating e.g. cervical

CC cancers, melanomas, colorectal adenocarcinomas, Wilms' tumour, sarcomas,

CC retinoblastoma, myosarcomas, lung carcinomas, leukemias, such as chronic

CC myelogenous leukemia, promyelocytic leukemia, monocytic leukemia, and

CC myeloid leukemia, and lymphomas such as histiocytic lymphoma, anhydric

CC hereditary ectodermal dysplasia, congenital alveolar dysplasia,

CC epithelial dysplasia of the cervix, fibrous dysplasia of bone, and

CC mammary dysplasia, hyperplasias, e.g. endometrial, adrenal, breast,

CC prostate or thyroid hyperplasias or pseudoepitheliomatous hyperplasia of

CC the skin.

XX Sequence 748 BP; 198 A; 152 C; 171 G; 211 T; 16 other;

Query Match 24.1%; Score 67; DB 20; Length 748;

Best Local Similarity 98.5%; Pred. No. 2.2e-07;

Matches 67; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 211 GCGGACCCATCGGAGCGCTAACCTGGATCTCCGAGGCGCTGGCGGAGCGCGCCACCTGGN 270

DB 71 GCGGACCCATCGGAGCGCTAACCTGGATCTCCGAGGCGCTGGCGGAGCGCGCCACCTGGG 130

QY 271 GGGGCATT 278

DB 131 GGGGCATT 138

RESULT 15

AAZ15699

XX AAZ15699 standard; cDNA; 754 BP.

XX AAZ15699;

XX 12-OCT-1999 (first entry)

XX Human gene expression product cDNA sequence SEQ ID NO:3168.

XX Human; gene; gene expression product; diagnosis; therapy; probe;

XX detection; mapping; tissue typing; profiling; forensic; cancer;

XX genetic analysis; colorectal cancer; breast cancer; lung cancer; ss.

XX Homo sapiens.

XX W09938972-A2.

XX 05-AUG-1999.

XX 28-JAN-1999; 99W0-US01619.

XX 03-APR-1998; 98US-0080666.

XX 28-JAN-1998; 98US-0072910.

PR 24-FEB-1998; 98US-0075954.

PR 31-MAR-1998; 98US-0080114.

PR 03-APR-1998; 98US-0080515.

XX (CHIR) CHIRON CORP.

PA (HYSE-) HYSEQ INC.

XX Crkvenjakov R, Dickson M, Drmanac R, Drmanac S;

PI Escobedo J, Garcia PD, Garcia V, Giese K, Innis MA;

PI Jones WL, Kassam A, Kennedy GC, Kita D, Labat I;

PI Lanson G, Leshkowitz D, Pot D, Randazzo F, Reinhard C;

PI Stache-Crain B, Sudduth-Klinger J, Williams LT;

XX WPI; 1999-49402/41.

XX Novel human genes and their expression products which are

PT differentially expressed in different cell types

PT Claim 1; Page 1522; 2479pp; English.

CC The present invention describes a library of human polynucleotides

CC comprising the sequences given in AAZ12532 to AAZ17779. Also described is

CC a method of detecting differentially expressed genes correlated with the

CC cancerous state of a mammalian cell, comprising detecting at least one

CC differentially expressed gene product in a test sample from a cell

CC suspected of being cancerous, where the gene product is encoded by one

CC of the 5248 polynucleotide sequences given in AAZ12532 to AAZ17779. The

CC polynucleotides can be used as a source of primers and probes, which can

CC be used for a variety of purpose, e.g. detection of expression levels,

CC mapping, tissue typing or profiling, forensics, genetic analysis and

CC detection of polymorphisms. Polypeptides encoded by the polynucleotides

CC can be used for raising antibodies for experimental, diagnostic and

CC therapeutic purposes. The polynucleotides may also be used to construct

CC arrays for diagnostics (which may be used to determine function of an

CC encoded protein); and to detect differences in expression levels between

CC two cells (e.g. to identify abnormal or diseased tissue in a human, to

CC identify a genetic predisposition or susceptibility to a disease such as

CC cancer). The polynucleotides of the invention are especially used in the

CC diagnosis, prognosis and management of colorectal cancer, breast cancer,

CC and lung cancer. The polynucleotides can also be used to screen for

CC peptide analogues and antagonists.

XX Sequence 754 BP; 199 A; 150 C; 170 G; 211 T; 24 other;

Query Match 24.1%; Score 67; DB 20; Length 754;

Best Local Similarity 98.5%; Pred. No. 2.2e-07;

Matches 67; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 211 GCGGACCCATCGGAGCGCTAACCTGGATCTCCGAGGCGCTGGCGGAGCGCGCCACCTGGN 270

DB 75 GCGGACCCATCGGAGCGCTAACCTGGATCTCCGAGGCGCTGGCGGAGCGCGCCACCTGGA 134

QY 271 GGGGCATT 278

DB 135 GGGGCATT 142

Search completed: August 27, 2003, 01:18:45

Job time : 1471 secs

Pred. No. is the number of results predicted by chance to have a

RESULT 1	DEFINITION	REFERENCE	TITLE
AF174602	ACCESSION	AUTHOR	
LOCUS	VERSION		
	KEYWORDS		
	SOURCE		
	ORGANISM		

ALIGNMENTS

277 bp DNA linear PRI 31-OCT-1999
n Fbx22 (FBX22) gene, partial cds.

AF174602.1 GI:6164746

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa;

Mammalia; Eutheria; 1997; 44; 377;

I (bases 1 to 277)

Cenciarelli, C., Chia
and Pagano M

and PAYANO, M.
Identification of a

3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842

JOURNAL MEDLINE 20003060
 PUBMED 10531035
 REFERENCE 2 (bases 1 to 277)
 AUTHORS Chiau, D.S. and Pagano, M.
 TITLE Direct Submission
 JOURNAL Submitted (03-AUG-1999) Pathology, NYU Medical Center, 550 First Ave. MSB 548, New York, NY 10016
 FEATURES Location/Qualifiers
 source 1..277
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 gene <34..>277
 /gene="FBX22"
 mRNA <34..>277
 /gene="FBX22"
 CDS 34..>277
 /product="F-box protein Fbx22"
 /note="FBX22"
 /note="contains one F-box"
 /codon_start=1
 /product="F-box protein Fbx22"
 /protein_id="AAF04523.1"
 /db_xref="GI:6164747"
 /translation="MEPAGAGCGGSSVDPSTFVLSNLAEEVVERVLTFLPAKALLRV ACVCLRWECVRRVLRTHRSVTWISAGLAERAGHLEGH"
 BASE COUNT 36 a 78 c 108 g 55 t
 ORIGIN
 Query Match 86.9%; Score 241.6; DB 9; Length 277;
 Best Local Similarity 95.7%; Pred. No. 5.8e-41;
 Matches 267; Conservative 0; Mismatches 9; Indels 3; Gaps 2;
 QY 1 CCGTAGTACTGGNTCCCGGGGGCTGGTAGAATGGAGCCGGTAGTGGTGGCGCGAG 60
 Db 1 CCGTAGTACTGGCTCCCGGGGGCTGGTAGAATGGAGCCGGTAGTGGTGGCGCGAG 60
 QY 61 TCCCGGGNTCCCTCCGTACACCGCGGA-NACCTTCGTGTGAGTAACCTGGCGGAGTGG 119
 Db 59 GTCCCGGCTCTCCGTACACCGCGGAGCACCTTCGTGTGAGTAACCTGGCGGAGTGG 118
 QY 120 TGGAGCGTGTGCTACCTTCCTGCCCGCCAAAGGGGTTGCTCGGGGTGGCCCTGCGTGC 179
 Db 119 TGGAGCGTGTGCTACCTTCCTGCCCGCCAAAGGGGTTGCTCGGGGTGGCCCTGCGTGC 178
 QY 180 GCTTATGAGGAGTGTGTCGCCAGAGTATTGGCGACCCATCGGAGCGTAACCTGGATCT 239
 Db 179 GCTTATGAGGAGTGTGTCGCCAGAGTATTGGCGACCCATCGGAGCGTAACCTGGATCT 238
 QY 240 CCGCAGGCGTGGCGAGGCGGCCACCTCGGNGGGGCATT 278
 Db 239 CCGCAGGCGTGGCGAGGCGGCCACCTCGGNGGGGCATT 277
 RESULT 2
 BD150672
 LOCUS BD150672 792 bp DNA linear PAT 17-JAN-2003
 DEFINITION Primer for synthesizing full-length cDNA and use thereof.
 ACCESSION BD150672
 VERSION BD150672.1 GI:27856430
 KEYWORDS JP 2002191363-A/5515.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 Oca, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J., Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.
 TITLE Primer for synthesizing full-length cDNA and use thereof
 JOURNAL Patent: JP 2002191363-A 5515 09-JUL-2002;
 COMMENT HELIX RESEARCH INSTITUTE
 OS Homo sapiens (human)
 PN JP 2002191363-A/5515
 PD 09-JUL-2002
 PF 28-JUL-2000 JP 2000280990
 PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU SAITO
 PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU, KEIICHI NAGAI, TETSUJI OTSUKI
 PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12P21/02, C12Q1/68, C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
 Primer for synthesizing full-length cDNA and use thereof FH Key
 Location/Qualifiers
 FT source 1..792
 FT /organism="Homo sapiens"
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 BASE COUNT 183 a 180 c 235 g 191 t 3 others
 ORIGIN
 Query Match 83.5%; Score 232; DB 6; Length 792;
 Best Local Similarity 95.5%; Pred. No. 5e-39;
 Matches 257; Conservative 0; Mismatches 10; Indels 2; Gaps 2;
 QY 12 GNTCCGGCGGCTGGTAGAATGGAGCCGGTAGTGGTGGCGGAGTCCCG-GGNTC 70
 Db 65 GTTCCGGCGGCTGGTAGAATGGAGCCGGTAGTGGTGGCGGAGTCCCGCGGCTC 124
 QY 71 CTCGTAGACCCGCGGA-NACCTTCGTGTGAGTAACCTGGCGGAGTGGTGGAGCGTGT 129
 Db 125 CTCGTAGACCCGCGGAGCACCTTCGTGTGAGTAACCTGGCGGAGTGGTGGAGCGTGT 184
 QY 130 GCTACCTTCCTGCCCGCCAAAGGGTTCCTCGGGGTGGCCCTGCGTGCCTCTTATGGAG 189
 Db 185 GCTACCTTCCTGCCCGCCAAAGGGTTCCTCGGGGTGGCCCTGCGTGCCTCTTATGGAG 244
 QY 190 GGAGTGTGTGCGCAGAGTATTGGCGACCCATCGGAGCGTAACCTGGATCTCCGAGGCGCT 249
 Db 245 GGAGTGTGTGCGCAGAGTATTGGCGACCCATCGGAGCGTAACCTGGATCTCCGAGGCGCT 304
 QY 250 GCGGAGGCGGCGCCACCTCGGNGGGGCATT 278
 Db 305 GCGGAGGCGGCGCCACCTCGGAGGCGCAT 333
 RESULT 3
 AX405640
 LOCUS AX405640 1298 bp DNA linear PAT 14-JUN-2002
 DEFINITION Sequence 55 from Patent WO022660.
 ACCESSION AX405640
 VERSION AX405640.1 GI:21438719
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 Tang, Y.T., Liu, C., Zhou, P., Asundi, V., Zhang, J., Zhao, Q.A., Ren, F., Xue, A.J., Yang, Y., Wehrman, T. and Drmanac, R.T.
 TITLE Novel nucleic acids and polypeptides
 JOURNAL Patent: WO 022660-A 55 21-MAR-2002;
 HYSEQ, INC. (US)
 FEATURES Location/Qualifiers
 source 1..1298
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 48..1259
 /note="unnamed protein product"
 /codon_start=1
 CDS

```

/protein_id="CAD34747.1"
/db_xref="GI:21438720"
/translation="MBPVCGCGRGSSVDRSTFVLSNLAEEVVERVLTLPKALLR
VACVRLRECVRLVTRHSVTWISAGLAEGHCLVRVVAEELENVRLPHVY
LYMADSETFISLSECHGRKARKRTSMETALALEKLPKQCQVLGIVTGIVPMGS
GSRNPQIEIGESGALLFPQIGIKIOPHFIFKDPKNFLERHQITEVGLLDNPELR
VVLVFGNCKVGSANVLOQVSTFSDMNIILAGQVDNLSLTSKPNLIDIASGVV
GLFSGHRIGASVAVLNEDVDEKTAEAAMRLKAANIPEHNTIGFMFACVGRGFQY
RAGKNVEADAFRRFFSVPLGFFNGEIGCDRIVTGNFILRKNEVKDDDLFHSYTT
IMALIHGSSK"
BASE COUNT      337 a      268 c      353 g      340 t
ORIGIN

```

```

Query Match      83.5%; Score 232; DB 6; Length 1298;
Best Local Similarity 95.5%; Pred. No. 4.5e-39;
Matches 257; Conservative 0; Mismatches 10; Indels 2; Gaps 2;

```

```

QY 12 GNTTCCGGCGGCTGTGAGAAATGGAGCGGTAGTGTGCGGCGAGTCCCG-GGNTC 70
D 12 GNTTCCGGCGGCTGTGAGAAATGGAGCGGTAGTGTGCGGCGAGTCCCG-GGNTC 70
Db 26 GGTTCGGCGGCGGCTGTGAGAAATGGAGCGGTAGTGTGCGGCGAGTCCCGGCGCTC 85
QY 71 CTCGTAGACCCCGCGGA-NACCTTCGTTGAGTAACCTGCGGAGGTGCTGGAGCGTGT 129
D 71 CTCGTAGACCCCGCGGA-NACCTTCGTTGAGTAACCTGCGGAGGTGCTGGAGCGTGT 129
Db 86 CTCGTAGACCCCGCGGAGCACCCTTCGTTGAGTAACCTGCGGAGGTGCTGGAGCGTGT 145
QY 130 GCTCACCTTCTGCGCGCCCAAGCGTTGCTGCGGCGTGCCTGCGGCGCTTATGGAG 189
D 130 GCTCACCTTCTGCGCGCCCAAGCGTTGCTGCGGCGTGCCTGCGGCGCTTATGGAG 189
Db 146 GCTCACCTTCTGCGCGCCCAAGCGTTGCTGCGGCGTGCCTGCGGCGCTTATGGAG 205
QY 190 GGAGTGTGTCGCGAGAGTATTGCGGACCCATCGGAGCGTAACCTGGATCTCCGAGCGCT 249
D 190 GGAGTGTGTCGCGAGAGTATTGCGGACCCATCGGAGCGTAACCTGGATCTCCGAGCGCT 249
Db 206 GGAGTGTGTCGCGAGAGTATTGCGGACCCATCGGAGCGTAACCTGGATCTCCGAGCGCT 265
QY 250 GCGGAGCGCGGCGCACCTGCGGGGCATT 278
D 250 GCGGAGCGCGGCGCACCTGCGGGGCATT 278
Db 266 GCGGAGCGCGGCGCACCTGCGGGGCATT 294

```

```

RESULT 4
BD158807
LOCUS      BD158807      1486 bp      DNA      linear      PAT 17-JAN-2003
DEFINITION      Primer for synthesizing full-length cDNA and use thereof.
ACCESSION      BD158807
VERSION      BD158807.1 GI:27864565
KEYWORDS      JP 2002191363-A/13650.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Ota,T., Isogai,T., Nishikawa,T., Sugiyama,T., Saito,K., Yamamoto,J.,
Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
Primer for synthesizing full-length cDNA and use thereof
Patent: JP 2002191363-A 13650 09-JUL-2002;
HELIX RESEARCH INSTITUTE
OS      Homo sapiens (human)
PN      JP 2002191363-A/13650
PD      09-JUL-2002
PF      28-JUL-2000 JP 2002080990
PI      TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU
SAITO,
PI      JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
PI      KEIICHI NAGAI, TETSUJI OTSUKI
PC
C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
10, C12P21/02, C12P1/68/ C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH key

```

```

FT CDS      Location/Qualifiers
           1..1486
           /organism="Homo sapiens"
FEATURES
source

```

```

/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT      369 a      301 c      367 g      449 t
ORIGIN
Query Match      83.5%; Score 232; DB 6; Length 1486;
Best Local Similarity 95.5%; Pred. No. 4.4e-39;
Matches 257; Conservative 0; Mismatches 10; Indels 2; Gaps 2;
QY 12 GNTTCCGGCGGCTGTGAGAAATGGAGCGGTAGTGTGCGGCGAGTCCCG-GGNTC 70
D 12 GNTTCCGGCGGCTGTGAGAAATGGAGCGGTAGTGTGCGGCGAGTCCCG-GGNTC 70
Db 65 GGTTCGGCGGCGGCTGTGAGAAATGGAGCGGTAGTGTGCGGCGAGTCCCGGCGCTC 124
QY 71 CTCGTAGACCCCGCGGA-NACCTTCGTTGAGTAACCTGCGGAGGTGCTGGAGCGTGT 129
D 71 CTCGTAGACCCCGCGGA-NACCTTCGTTGAGTAACCTGCGGAGGTGCTGGAGCGTGT 129
Db 125 CTCGTAGACCCCGCGGAGCACCCTTCGTTGAGTAACCTGCGGAGGTGCTGGAGCGTGT 184
QY 130 GCTCACCTTCTGCGCGCCCAAGCGTTGCTGCGGCGTGCCTGCGGCGCTTATGGAG 189
D 130 GCTCACCTTCTGCGCGCCCAAGCGTTGCTGCGGCGTGCCTGCGGCGCTTATGGAG 189
Db 185 GCTCACCTTCTGCGCGCCCAAGCGTTGCTGCGGCGTGCCTGCGGCGCTTATGGAG 244
QY 190 GGAGTGTGTCGCGAGAGTATTGCGGACCCATCGGAGCGTAACCTGGATCTCCGAGCGCT 249
D 190 GGAGTGTGTCGCGAGAGTATTGCGGACCCATCGGAGCGTAACCTGGATCTCCGAGCGCT 249
Db 245 GGAGTGTGTCGCGAGAGTATTGCGGACCCATCGGAGCGTAACCTGGATCTCCGAGCGCT 304
QY 250 GCGGAGCGCGGCGCACCTGCGGGGCATT 278
D 250 GCGGAGCGCGGCGCACCTGCGGGGCATT 278
Db 305 GCGGAGCGCGGCGCACCTGCGGGGCATT 333

```

```

RESULT 5
AK024048
LOCUS      AK024048      1486 bp      mRNA      linear      PRI 01-AUG-2002
DEFINITION      Homo sapiens cDNA FLJ13986 fis, clone Y79AA100193, weakly similar
to Homo sapiens F-box protein Fbx22 (FBX22) gene.
ACCESSION      AK024048
VERSION      AK024048.1 GI:10436316
KEYWORDS      oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
Nishikawa,T., Nagai,K., Sugano,S., Shiratori,A., Sudo,H.,
Wagatsuma,M., Hosoi,T., Kaku,Y., Kodaira,H., Kondo,H.,
Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K.,
Ono,Y., Takiguchi,S., Watanabe,S., Kimura,K., Murakami,K.,
Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A.,
Nakamura,Y., Nagahara,K., Masuho,Y., Ninomiya,K. and Iwayanagi,T.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 1486)
Isogai,T. and Otsuki,T.
Direct Submission
Submitted (23-AUG-2000) Takao Isogai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel.81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing. Research Association for Biotechnology; cDNA library
construction, 5' & 3'-end one pass sequencing and clone selection:
Helix Research Institute (supported by Japan Key Technology Center
etc.) and Department of Virology, Institute of Medical Science,
University of Tokyo.
Location/Qualifiers
1..1486
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="Y79AA100193"
/cell_line="Y79"
/cell_type="retinoblastoma"
FEATURES
source

```


/clone_lib="Y79Aa1"
/note="cloning vector: pME18SFL3"
87. .917
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB14798.1"
/db_xref="GI:10436317"

FEATURES

source

Location/Qualifiers

1. .2335

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:48729 IMAGE:5246992"

/tissue_type="Brain, fetal, whole pooled"

/clone_lib="NIH_MGC_121"

/lab_host="DH10B"

/note="vector: pCMV-SPORT6"

73. .1284

/codon_start=1

/product="F-box only protein 22"

/protein_id="AAH41691.1"

/db_xref="GI:27469585"

/db_xref="LocusID:26263"

/translation="MEPVGCGGCGSSVDPSTFVLSNLAEEVVERVLTFLPAKALLR
VACVRLWRECVRLVTRSVTWISAGLAEGHLECHCLRVVVAELENRLPHTV
LYMADSETFISLEECRHKRARKTSMETALALEKLPKQCQVLGIVTPGIVTPMGS
GSRNPQIEIGESGFALLFQIEGKIQPFHFIDKPNLTLERHQLTEVGLLONPELR
VVLVFGNCKVGSNLYQQVSTFSDMNILAGQVDNLSLSTSEKYNPLDIDASGVV
GLSFNGHRIOQATVLLNEDVDEKTAEMORLKAANIPEHNTIGFMFACVGRGFOY
RAGSNVEADAFKFFPSVPLGFFGNGEIGCDRIVTGNFILRKCNKVEKDDDLFHSYTT
IMALIHLSGSK"

BASE COUNT 369 a 301 c 367 g 449 t

ORIGIN

Query Match 83.5%; Score 232; DB 9; Length 1486;

Best Local Similarity 95.5%; Pred. No. 4.4e-39;

Matches 257; Conservative 0; Mismatches 10; Indels 2; Gaps 2;

Qy 12 GNTTCCGCGCGGCTGGTGAGGAATGGAGCGGTAGNTGCTTGGCGGCGAGTCCCG-GGNTC 70

Db 65 GGTTCGCGCGGCTGGTGAGGAATGGAGCGGTAGGCTGTGCGCGGAGTCCCGCGCTC 124

Qy 71 CTCGCTAGACCCGCGGA-NACCTTCGTTGTGAGTAACCTGCGGAGGTGGTGAGCGCTGT 129

Db 125 CTCGCTAGACCCGCGGAGCACCTTCGTTGTGAGTAACCTGCGGAGGTGGTGAGCGCTGT 184

Qy 130 GCTCACCTTCCTGCCCGCGGAGCGGTGCTCCGGGTGCGCTGCTGCGCGCTTATGGAG 189

Db 185 GCTCACCTTCCTGCCCGCGGAGCGGTGCTCCGGGTGCGCTGCTGCGCGCTTATGGAG 244

Qy 190 GGAGTGTGTGCGGAGAGTATTGCGGACCCATCGGAGCGTAACCTGGATCTCCCGAGCGCT 249

Db 245 GGAGTGTGTGCGGAGAGTATTGCGGACCCATCGGAGCGTAACCTGGATCTCCCGAGCGCT 304

Qy 250 GCGGAGCGCGGCGCACCTGNGGGGCATT 278

Db 305 GCGGAGCGCGGCGCACCTGNGGGGCATT 333

RESULT 6

LOCUS BC041691

DEFINITION Homo sapiens, F-box only protein 22, clone MGC:48729 IMAGE:5246992,
mRNA, complete cds.

ACCESSION BC041691

VERSION BC041691.1 GI:27469584

KEYWORDS MGC.

SOURCE Homo sapiens

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 2335)
Strausberg, R.
Direct Submission
Submitted (23-DEC-2002), National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
contact: amadan@systemsbiology.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha
Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 84 Row: e Column: 17
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 1342999.

FEATURES

source

Location/Qualifiers

1. .2335

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:48729 IMAGE:5246992"

/tissue_type="Brain, fetal, whole pooled"

/clone_lib="NIH_MGC_121"

/lab_host="DH10B"

/note="vector: pCMV-SPORT6"

73. .1284

/codon_start=1

/product="F-box only protein 22"

/protein_id="AAH41691.1"

/db_xref="GI:27469585"

/db_xref="LocusID:26263"

/translation="MEPVGCGGCGSSVDPSTFVLSNLAEEVVERVLTFLPAKALLR
VACVRLWRECVRLVTRSVTWISAGLAEGHLECHCLRVVVAELENRLPHTV
LYMADSETFISLEECRHKRARKTSMETALALEKLPKQCQVLGIVTPGIVTPMGS
GSRNPQIEIGESGFALLFQIEGKIQPFHFIDKPNLTLERHQLTEVGLLONPELR
VVLVFGNCKVGSNLYQQVSTFSDMNILAGQVDNLSLSTSEKYNPLDIDASGVV
GLSFNGHRIOQATVLLNEDVDEKTAEMORLKAANIPEHNTIGFMFACVGRGFOY
RAGSNVEADAFKFFPSVPLGFFGNGEIGCDRIVTGNFILRKCNKVEKDDDLFHSYTT
IMALIHLSGSK"

BASE COUNT 588 a 426 c 523 g 698 t

ORIGIN

Query Match 83.5%; Score 232; DB 9; Length 2335;

Best Local Similarity 95.5%; Pred. No. 4e-39;

Matches 257; Conservative 0; Mismatches 10; Indels 2; Gaps 2;

Qy 12 GNTTCCGCGCGGCTGGTGAGGAATGGAGCGGTAGNTGCTTGGCGGCGAGTCCCG-GGNTC 70

Db 51 GGTTCGCGCGGCTGGTGAGGAATGGAGCGGTAGGCTGTGCGCGGAGTCCCGCGCTC 110

Qy 71 CTCGCTAGACCCGCGGA-NACCTTCGTTGTGAGTAACCTGCGGAGGTGGTGAGCGCTGT 129

Db 111 CTCGCTAGACCCGCGGAGCACCTTCGTTGTGAGTAACCTGCGGAGGTGGTGAGCGCTGT 170

Qy 130 GCTCACCTTCCTGCCCGCGGAGCGGTGCTCCGGGTGCGCTGCTGCGCGCTTATGGAG 189

Db 171 GCTCACCTTCCTGCCCGCGGAGCGGTGCTCCGGGTGCGCTGCTGCGCGCTTATGGAG 230

Qy 190 GGAGTGTGTGCGGAGAGTATTGCGGACCCATCGGAGCGTAACCTGGATCTCCCGAGCGCT 249

Db 231 GGAGTGTGTGCGGAGAGTATTGCGGACCCATCGGAGCGTAACCTGGATCTCCCGAGCGCT 290

Qy 250 GCGGAGCGCGGCGCACCTGNGGGGCATT 278

Db 291 GCGGAGCGCGGCGCACCTGNGGGGCATT 319

RESULT 7

LOCUS AY005144

DEFINITION Homo sapiens F-box protein FBX22p44 mRNA, complete cds.

ACCESSION AY005144

VERSION AY005144.1 GI:22073861

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1212)
Tan, P. and Pan, Z.-Q.
FBX22p44: a novel human F-box protein predominantly expressed in
the liver
Unpublished
JOURNAL REFERENCE 2 (bases 1 to 1212)

McKernan, X.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Vallalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettanan, M., Madan, A., Rodrigues, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shrivchenko, Y.,
Bouffard, J.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzyzanski, M.I., Skalska, U., Smalls, D.E.,
Schnierch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
12477932
2 (bases 1 to 2057)
Strausberg, R.
Direct Submission
Submitted (03-DEC-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NTH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgmpbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulsegod, H.,
Kowis, C.H., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IHAK Plate: 31 Row: h Column: 24
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, GenomeScan gene prediction, Similarity but not identity
to protein.
Location/Qualifiers
1. .2057
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="MGC:25463 IMAGE:4458296"
/tissue_type="Mammary tumor. Metallothionien-TGF alpha
model, 10 month old virgin mouse. Taken by biopsy."
/clone_lib="NCI CGAP_Mam1"
/lab_host="DH10B"
/note="vector: pCMV-SPORT6"
1. .2057
/codon_start=1
/product="F-box only protein 22"
/protein_id="AAH18273.1"
/db_xref="GI:17390644"
/db_xref="LocusID:711999"
/translation="MADETFISLEECRGHRRARKRTTETACALEKLPKQCVLGI
VTPGVITPMGSGNRPOEITGSGFALLPFGIQTPHFTKDSKNLTIERHQL
TEVGLDNPPELRVVLFCYCNKVCASNYLHRVSTFSDMTILLAGOVNLSLTCE
KNPLDIDATGVGLSFSGHRQTSATVLTEDVNDAKTVEAAMQRKAKANIPQNTIGF
MFACVGRGFTYVRAKGNVEADRAKFRFFVSPVLPFGFGGICGDRIVTGNFLRRKNC
VKEDFLFSYTTINALVHLGTSK"
454 c 541 g 522 t

BC039024	2188 bp	mRNA	linear	PRI 06-NOV-2002
LOCUS				
DEFINITION	Homo sapiens, similar to RIKEN CDNA 1600016C16 gene, clone MGC:47575 IMAGE:6071245, mRNA, complete cds.			
ACCESSION	BC039024			
VERSION	BC039024.1			
KEYWORDS	GI:24659488			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
AUTHORS	1 (bases 1 to 2188)			
TITLE	Strausberg, R.			
JOURNAL	Direct Submission			
	Submitted (01-NOV-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA			
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov			
COMMENT	Contact: MGC help desk Email: cgabbs@mail.nih.gov Tissue Procurement: DCTD/DTP/Gazdar CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arranged by: The I.M.A.G.E. Consortium (LNL)			

```

Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc\_mgc@hri.nih.gov
Akhter N., Ayele K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
Maduro, Q.L., Mastello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
Tsurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
through the I.W.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 82 Row: p Column: 3
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 13442999.
Location/Qualifiers
1. 2188
/organism="Homo sapiens"
/mol_type="mRNA"
FEATURES
Source

```

```

CDS
/lab_host="DHIUB"
/note="Vector: pCHV-SPORT6"
237. .1136
/codon_start=1
/product="Similar to RIKEN CDNA 1600016C16 gene"
/protein_id="AAH39024.1"
/db_xref="GI:24653489"
/translation="MADSETTISLEECGRHKRARRKRTSMETALALEKLFPPKOCVLGI
VTPGIVPMGSGNRNQOEIGEGSFALLFPQIEGIKIQPHFTKDPKNLTLERHQL
TEVGLIDNPELRVLNCFYCNCKVGASNYLQGVSTFFSDMMIILAGGOVDNLSLSTSE
KNPDIADSGVVGLGSFSGRIQSTAVLLNEDVSDKETAAMORAKAANIPEHNTIGF
MKACVGRGFQYIRAKNGVNEADAFKRFPSVLPFGFGNGEIGCDIRIVTGCNFIKRCNE
VFCDLDEHSYITMALIHGSSK"
665 a 393 C 463 g 667 t
BASE COUNT
ORIGIN
Query Match 46.1%; Score 138.2; DB 9; Length 2188;
Best Local Similarity 91.7%; Pred. No. 2.9e-17;
Matches 155; Conservative 0; Mismatches 12; Indels 2; Gaps 2;
QY 12 GNTTCCGCGGGCTGGTGAGGAATGGACCGGTAGTCTTTCGCGCGAGTCCCG-GGNTC 70

```

Db 42 GGTCCGGCGGCTGTGAGGAATGGAGCGGTAGTGTCTGCGGAGTCCCGGCTC 101
 QY 71 CTCGGTAGACCCGGGA-NACCTTCGTGTTAGTAACCTGCGGAGTGTGGAGCGCTGT 129
 Db 102 CTCGGTAGACCCGGAGACCTTCGTGTTAGTAACCTGCGGAGTGTGGAGCGCTGT 161
 QY 130 GCTCACCTTCCTGTCGCCGCCAAGCGTGTCTGCGGTCGCCCTGCGCTGCC 178
 Db 162 GCTCACCTTCCTGTCGCCGCCAAGCGTGTCTGCGGTCGCCCTGCGCTGATGTC 210

RESULT 11

AC087456

LOCUS

DEFINITION

AC087456

ACCESSION

AC087456.3

KEYWORDS

HTG.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 170945)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,

Barna,N., Bastien,V., Boguslavskiy,L., Boukhgalter,B., Brown,A.,

Camarata,J., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S.,

Collamore,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S.,

Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J.,

Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N.,

Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,

Jones,C., Karatas,A., LaRocque,K., Lamazares,R., Landers,T.,

Lehoczky,J., Levine,R., Liu,G., MacLean,C., Macdonald,P.,

Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K.,

McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,

Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H.,

O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,

Phunkhang,P., Plerre,N., Pollara,V., Raymond,C., Retta,R.,

Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M.,

Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S., Severy,P.,

Sougnuez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,

Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,

Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A.,

Wilson,B., Wu,X., Wyman,D., Ye-W.J., Young,G., Zainoun,J.,

Zembek,L., Zimmer,A. and Zody,M.

Zembek,L., Zimmer,A. and Zody,M.

Zembek,L., Zimmer,A. and Zody,M.

Zembek,L., Zimmer,A. and Zody,M.

Zembek,L., Zimmer,A. and Zody,M.

Zembek,L., Zimmer,A. and Zody,M.

Zembek,L., Zimmer,A. and Zody,M.

Zembek,L., Zimmer,A. and Zody,M.

Zembek,L., Zimmer,A. and Zody,M.

Zembek,L., Zimmer,A. and Zody,M.

Zembek,L., Zimmer,A. and Zody,M.

Zembek,L., Zimmer,A. and Zody,M.

Zembek,L., Zimmer,A. and Zody,M.

Zembek,L., Zimmer,A. and Zody,M.

Zembek,L., Zimmer,A. and Zody,M.

Zembek,L., Zimmer,A. and Zody,M.

Zembek,L., Zimmer,A. and Zody,M.

Zembek,L., Zimmer,A. and Zody,M.

Zembek,L., Zimmer,A. and Zody,M.

Zembek,L., Zimmer,A. and Zody,M.

Zembek,L., Zimmer,A. and Zody,M.

Zembek,L., Zimmer,A. and Zody,M.

Zembek,L., Zimmer,A. and Zody,M.

Zembek,L., Zimmer,A. and Zody,M.

Zembek,L., Zimmer,A. and Zody,M.

Zembek,L., Zimmer,A. and Zody,M.

Zembek,L., Zimmer,A. and Zody,M.

Zembek,L., Zimmer,A. and Zody,M.

TITLE

JOURNAL

COMMENT

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence.submissions@genome.wi.mit.edu

----- Project Information

Center project name: L11871

Center clone name: 26_A13

FEATURES

source

1..170945

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="15"

/map="15"

/clone="RP11-26A13"

/clone_lib="RPC1-11 Human Male BAC"

/complement(1140..1434)

/rpt_family="L1MB7"

/complement(4552..4702)

/rpt_family="MIR"

5224..5454

/rpt_family="L1Med"

5460..5626

/rpt_family="L1Med"

5627..5923

/rpt_family="AluJb"

5924..6121

/rpt_family="L1Med"

6272..6920

/rpt_family="L1Med"

7331..7448

/rpt_family="AluSx"

7449..7473

/rpt_family="AluSx"

7474..7637

/rpt_family="AluSx"

7970..8242

/rpt_family="AluY"

complement(8243..8529)

/rpt_family="AluSp"

complement(8383..8397)

/note="probably CC; possibly CCC"

/note="<30 qual SINGL region"

complement(8392)

/note="probably T"

8620..9056

/rpt_family="L1MC/D"

complement(10270..10560)

/rpt_family="AluY"

complement(10730..10879)

/rpt_family="FAM"

complement(11424..11636)

/rpt_family="L1MA8"

12025..12327

/rpt_family="AluSx"

12328..12352

/rpt_family="(TA)n"

12426..12461

/rpt_family="AT-rich"

complement(14454..14496)

/rpt_family="MLTIL"

14773..15001

repeat_region

repeat_region

repeat_region

repeat_region

```

/rpt_family="Charlielb"
15463..15594
/rpt_family="FLAM_C"
15661..15853
/rpt_family="Charlielb"
complement(15977..16136)
/rpt_family="MER5A"
16519..16800
/rpt_family="AluJb"
complement(17874..18169)
/rpt_family="AluY"
complement(18486..19040)
/rpt_family="LIMC4"
complement(19043..19436)
/rpt_family="LIMC4a"
complement(19482..19654)
/rpt_family="LIMD3"
20205..20436
/rpt_family="MIR"
20663..21003
/rpt_family="AluYb8"
complement(21230..21308)
/rpt_family="MIR3"
complement(21309..21617)
/rpt_family="AluSx"
complement(21618..21648)
/rpt_family="MIR3"
complement(21655..21948)
/rpt_family="MER102b"
complement(22823..26183)
/rpt_family="L1PA7"
27802..28019
/rpt_family="MIR"
complement(28029..28318)
/rpt_family="AluSx"
28533..28668
/rpt_family="AluSx"
28669..28963
/rpt_family="AluSg"
28964..29125
/rpt_family="AluSx"
complement(29308..29648)
/rpt_family="MLT1B"
complement(29704..29847)
/rpt_family="L1MB4A"
complement(29959..30146)
/rpt_family="L1MC4"
31561..31998
/rpt_family="L1MA8"
complement(31999..32315)
/rpt_family="AluSx"
32316..32497
/rpt_family="L1MA8"
32541..32696
/rpt_family="MIR"
complement(32884..33060)
/rpt_family="MIR"
33589..33889
/rpt_family="AluJo"
complement(33974..34423)
/rpt_family="Tiggr3b"
34424..34718
/rpt_family="AluSx"
complement(34719..35309)
/rpt_family="Tiggr3b"
complement(35440..35606)
/rpt_family="L2"

```

Query Match 45.3%; Score 126; DB 9; Length 170945;
Best Local Similarity 93.2%; Pred. No. 3.5e-17;
Matches 151; Conservative 0; Mismatches 9; Indels 2; Gaps 2;
QY 12 GNTCCGGGGGCTGTGAGGAATGAGCGGTAGTGTTCGGCGAGTCCCG-GGNTC 70

```

Db 16986 GGTTCGGCGGGCTGTGAGGAATGAGCGGTAGTGTTCGGCGAGTCCCGGCTC 17045
QY 71 CTCCTGAGACCGCGGA-NACCTTCGTGTGAGTAACCTTGGCGGAGGTGTGTGAGGCTGT 129
Db 17046 CTCCTGAGACCGCGGAGCACCTTCGTGTGAGTAACCTTGGCGGAGGTGTGTGAGGCTGT 17105
QY 130 GCTCACCTTCCTGCGCCGCAAGCGTTCGTGCGGGTGGCGTG 171
Db 17106 GCTCACCTTCCTGCGCCGCAAGCGTTCGTGCGGGTGGCGTG 17147

RESULT 12
AC027104 180202 bp DNA linear PRI 02-MAR-2002
LOCUS Homo sapiens chromosome 15, clone RP11-326L17, complete sequence.
DEFINITION AC027104
ACCESSION AC027104.6 GI:18875285
VERSION HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 180202)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 15, clone RP11-326L17
Unpublished
2 (bases 1 to 180202)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
Galagan,J., Cardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczy,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
Meldrum,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Plerre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
Vassiliev,H., Viel,R., Vo.A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (26-MAR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 180202)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepe,Y., Colangelo,M., Collins,S., Collamore,A., Cook,A.,
Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Fato,S.,
Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
Landers,T., Lehoczy,J., Levine,R., Liu,G., MacLean,C.,
Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,
McEwan,P., McKernan,K., Meldrum,J., Meneus,L., Mihova,T.,
Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,
Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S.,
Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,

```

Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (25-FEB-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 4 (bases 1 to 180202)

TITLE JOURNAL REFERENCE AUTHORS

Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Collangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., DeAtrellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferraira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Glende, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, J., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., LaRoque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigglio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (02-MAR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Feb 25, 2002 this sequence version replaced gi:13489207.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

TITLE JOURNAL COMMENT

----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L8783
 Center clone name: 326_L17

FEATURES source

Location/Qualifiers
 1. .180202
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="15"
 /map="15"
 /clone="RP11-326L17"
 /clone_lib="RPC1-11 Human Male BAC"
 8. .426
 /rpt_family="LIMEC"
 2185. .2991
 /rpt_family="LIMEC"
 2992. .3288
 /rpt_family="AluX"
 3289. .3729
 /rpt_family="LIMEC"
 3730. .3991
 /rpt_family="AluJb"
 3999. .4221
 /rpt_family="LIMEC"
 4217. .4996
 /rpt_family="LIMB8"
 4997. .5123
 /rpt_family="FLAM_A"
 5149. .5644
 /rpt_family="LIMB8"
 5666. .6102
 /rpt_family="LIMEC"

repeat_region 6105. .6242
 /rpt_family="LIMEC"
 repeat_region 6259. .6559
 /rpt_family="AluX"
 repeat_region 6841. .6932
 /rpt_family="LIM4"
 repeat_region 7420. .7814
 /rpt_family="LIM4"
 repeat_region 7835. .8302
 /rpt_family="LIM4"
 repeat_region complement(8305. .8552)
 /rpt_family="AluX"
 repeat_region 8565. .8651
 /rpt_family="LIMB7"
 repeat_region complement(8652. .8948)
 /rpt_family="AluJo"
 repeat_region 8949. .9003
 /rpt_family="LIMB7"
 repeat_region 9023. .9114
 /rpt_family="LIMB5"
 repeat_region 9115. .9356
 /rpt_family="AluSg"
 repeat_region 9357. .9368
 /rpt_family="LIMB5"
 repeat_region 9369. .9410
 /rpt_family="(TAAA)n"
 repeat_region 9411. .10124
 /rpt_family="LIMB5"
 repeat_region complement(10125. .10710)
 /rpt_family="LTR43"
 repeat_region 10711. .11057
 /rpt_family="LIMB5"
 repeat_region 11058. .11372
 /rpt_family="AluY"
 repeat_region 11373. .11413
 /rpt_family="LIMB5"
 repeat_region 11414. .11922
 /rpt_family="MER44D"
 repeat_region 11923. .11983
 /rpt_family="LIMB5"
 repeat_region 11237. .12466
 /rpt_family="AluSg/x"
 repeat_region 12749. .12783
 /rpt_family="AT-rich"
 repeat_region complement(13393. .13516)
 /rpt_family="MIR"
 repeat_region complement(15210. .15352)
 /rpt_family="LIMC2"
 repeat_region complement(15353. .15415)
 /rpt_family="AluYb8"
 repeat_region complement(15416. .15696)
 /rpt_family="LIMC2"
 repeat_region 15723. .15743
 /rpt_family="AT-rich"
 repeat_region 15755. .15775
 /rpt_family="(T)n"
 repeat_region complement(15794. .18925)
 /rpt_family="LIMB5"
 repeat_region complement(19010. .19773)
 /rpt_family="LIMB5"
 repeat_region complement(20124. .20183)
 /rpt_family="LIMB5"
 repeat_region complement(20184. .20488)
 /rpt_family="AluY"
 repeat_region complement(20489. .20559)
 /rpt_family="LIMB5"
 repeat_region 20844. .21134
 /rpt_family="AluSp"

Query Match 45.3%; Score 126; DB 9; Length 180202;
 Best Local Similarity 93.2%; Pred. No. 3.5e-17;
 Matches 151; Conservative 0; Mismatches 9; Indels 2; Gaps 2;

Qy 12 GNTTCCGGCGGCTGGTGGAGAAATGAGCGGTAGTCTTGGCGGAGTCCCG-GGNTC 70
 Db 53774 GTTCCGGCGGCTGGTGGAGAAATGAGCGGTAGTCTTGGCGGAGTCCCGGCTC 53715
 Qy 71 GTCCTAGACCCCGCGGA-NACCTTCGTGTTGAGTAACTCGCGGAGTGGTGGAGCGCTG 129
 Db 53714 CTCCTAGACCCCGCGGACACTCGTGTGAGTAACTCGCGGAGTGGTGGAGCGCTG 53655
 Qy 130 GTCACCTCTCTCCCGCCCAAGCGGTGCTCGCGGTGGCGCTG 171
 Db 53654 GTCACCTCTCTCCCGCCCAAGCGGTGCTCGCGGTGGCGCTG 53613

RESULT 13

AC068838/c
 LOCUS AC068838 187649 bp DNA linear HTG 06-AUG-2002
 DEFINITION Homo sapiens chromosome 15 clone RP11-591G17 map 15, WORKING DRAFT
 AC068838
 VERSION HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 187649)

AUTHORS Birren,B., Nusbaum,C. and Lander,E.

TITLE Homo sapiens chromosome 15, clone RP11-591G17

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 187649)

AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
 Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
 Boguslavsky,L., Boukhgalter,B., Brown,A., Buckett,G.,
 Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
 Collymore,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S.,
 Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
 Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
 Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
 Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
 Levine,R., Liu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
 McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
 Meldrum,J., Meneus,L., Mihova,T., Miranda,C., Mienga,V., Morrow,J.,
 Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
 O'Neill,D., Olliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
 Pisan,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
 Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
 Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
 Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
 Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
 Young,G., Zainoun,J., Zimmer,A. and Zody,M.

Direct Submission
 Submitted (10-MAY-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE 3 (bases 1 to 187649)

AUTHORS Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
 Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B.,
 Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
 Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
 Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,
 Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
 Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
 Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
 Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
 McCarthy,M., Meldrum,J., Meneus,L., Mihova,T., Mienga,V.,
 Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
 O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
 Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
 Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
 Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
 Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
 Zembek,L., Zimmer,A. and Zody,M.

TITLE
JOURNAL

COMMENT

Direct Submission
 Submitted (06-AUG-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Aug 6, 2002 this sequence version replaced gi:13958576.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L7097

Center clone name: 591_G_17

----- Summary Statistics

Sequencing vector: M13; M7815; 4% of reads

Sequencing vector: Plasmid; n/a; 96% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 182998 bases at least Q40

Consensus quality: 185443 bases at least Q30

Consensus quality: 186227 bases at least Q20

Insert size: 188000; agarose-fp

Quality coverage: 5.3 in Q20 bases; agarose-fp

Quality coverage: 5.4 in Q20 ba

NOTE: This is a 'working draft' sequence. It currently

consists of 12 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.

This record will be updated with the finished sequence

as soon as it is available and the accession number will

be preserved.

1 2044: contig of 2044 bp in length

2045 2144: gap of 100 bp

2145 4878: contig of 2734 bp in length

4879 4978: gap of 100 bp

4979 48122: contig of 43144 bp in length

48123 48222: gap of 100 bp

48223 52851: contig of 4629 bp in length

52852 52951: gap of 100 bp

52952 58586: contig of 5635 bp in length

58587 58686: gap of 100 bp

58687 66926: contig of 8240 bp in length

66927 67026: gap of 100 bp

67027 82230: contig of 15204 bp in length

82231 82330: gap of 100 bp

82331 96640: contig of 14310 bp in length

96641 96740: gap of 100 bp

96741 122515: contig of 25775 bp in length

122516 122615: gap of 100 bp

122616 150859: contig of 28244 bp in length

150860 150959: gap of 100 bp

150960 185696: contig of 34737 bp in length

185697 185796: gap of 100 bp

185797 187649: contig of 1853 bp in length.

Location/Qualifiers

1. 187649

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="15"

/map="15"

/clone="RP11-591G17"

/glone_lib="RPC1-11 Human Male BAC"

1. 2044

/note="assembly_fragment"

clone_end:SP6

vector_side:left"

2145. 1878

/note="assembly_fragment"

FEATURES
source

misc_feature

misc_feature

```

misc_feature 4979. 48122
/Note="assembly_fragment"
misc_feature 4823. 52851
/Note="assembly_fragment"
misc_feature 5295. 58586
/Note="assembly_fragment"
misc_feature 5867. 66926
/Note="assembly_fragment"
misc_feature 6702. 82230
/Note="assembly_fragment"
misc_feature 8231. 96640
/Note="assembly_fragment"
misc_feature 9671. 122515
/Note="assembly_fragment"
misc_feature 12261. 150859
/Note="assembly_fragment"
misc_feature 15096. 185696
/Note="assembly_fragment"
misc_feature 18579. 187649
/Note="assembly_fragment"
vector_side:right
clone_end:T7
BASE COUNT 59871 a 36165 c 35300 g 55210 t 1103 others
ORIGIN
Query Match 45.3%; Score 126; DB 2; Length 187649;
Best Local Similarity 93.2%; Pred. No. 3.5e-17;
Matches 151; Conservative 0; Mismatches 9; Indels 2; Gaps 2;

QY 12 GNTTCCGGCGGCTGTGAGATGGAGCGGTAGTGTCTGCGCGGAGTCCCG-GGNTC 70
Db 142255 GTTTCCTGGCGGCTGTGAGATGGAGCGGTAGTGTCTGCGCGGAGTCCCGGCTC 142196

QY 71 CTCCTGTAGACCCGCGGA-NACCTTCGTGTTGAGTAACCTGCGCGAGGTGGAGCGCTGT 129
Db 142195 CTCCTGTAGACCCGCGGACCTTCGTGTTGAGTAACCTGCGCGAGGTGGAGCGCTGT 142136

QY 130 GCTCACCTTCTGCCCCCAAGGCGTTCCTGCGGTGGCCCTG 171
Db 142135 GCTCACCTTCTGCCCCCAAGGCGTTCCTGCGGTGGCCCTG 142094

RESULT 14
AC107831 207221 bp DNA linear HTG 17-APR-2002
LOCUS Mus musculus clone RP23-230P11, WORKING DRAFT SEQUENCE, 13 ordered
DEFINITION pieces.
AC107831
VERSION AC107831.3 GI:20163121
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 207221)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgaiter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepe,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliiev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
Landers,T., Lehoczy,J., Levine,R., Liu,G., MacLean,C.,
Macdonald,P., McKernan,K., Meldrim,J., Meneus,C.,
McEwan,P., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission
Submitted (17-APR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 17, 2002 this sequence version replaced gi:20148016.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center code: WIDR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L20581
Center clone name: 230_P11
----- Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 203635 bases at least Q40
Consensus quality: 205071 bases at least Q30
Consensus quality: 205592 bases at least Q20
Insert size: 200000; agarose-ef
Insert size: 206021; sum-of-contigs
Quality coverage: 7.8 in Q20 bases; agarose-ef
Quality coverage: 7.6 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 13 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as runs of N. The order of the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 1325: contig of 1325 bp in length
* 1326 1425: gap of 100 bp
* 1426 4349: contig of 2924 bp in length

```

```

Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S.,
Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission
Submitted (24-JAN-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 207221)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L.,
Boukhgaiter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepe,Y., Colangelo,M., Collins,S., Collymore,A.,
Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliiev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,
MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,N., McEwan,P., McKernan,K., Meldrim,J., Meneus,C.,
Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission
Submitted (17-APR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 17, 2002 this sequence version replaced gi:20148016.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center code: WIDR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L20581
Center clone name: 230_P11
----- Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 203635 bases at least Q40
Consensus quality: 205071 bases at least Q30
Consensus quality: 205592 bases at least Q20
Insert size: 200000; agarose-ef
Insert size: 206021; sum-of-contigs
Quality coverage: 7.8 in Q20 bases; agarose-ef
Quality coverage: 7.6 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 13 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as runs of N. The order of the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 1325: contig of 1325 bp in length
* 1326 1425: gap of 100 bp
* 1426 4349: contig of 2924 bp in length

```


REFERENCE
AUTHORS
TITLE
JOURNAL

2 (bases 1 to 252612)
Worley, K.C.

Direct Submission

Submitted (19-FEB-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

REFERENCE
AUTHORS
TITLE
JOURNAL

3 (bases 1 to 252612)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (13-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

COMMENT

On May 13, 2003 this sequence version replaced gi:24942379.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GLXJ

Center clone name: CH230-40C24

----- Summary Statistics

Assembly program: Atlas 3.0;

Consensus quality: 242384 bases at least Q40

Consensus quality: 244789 bases at least Q30

Consensus quality: 246474 bases at least Q20

Estimated insert size: 255317; sum-of-contigs estimation

Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)
* NOTE: This sequence may represent more than one clone.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 251112: contig of 251112 bp in length

* 251113 251212: gap of unknown length

* 251213 252612: contig of 1400 bp in length.

FEATURES

source

1. .252612

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

/clone="CH230-40C24"

247045..248261

/note="wgs_contig"

249152..251112

/note="wgs_contig"

BASE COUNT 67901 a 54137 c 56139 g 70169 t 4266 others

ORIGIN

Query Match

Best Local Similarity 29.9%; Score 83; DB 2; Length 252612;

Matches 92; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy

171 GCCTGTCCTTATGGAGGGAGTGTGTGGCAGAGTATTGCGGACCATCGGAGCGCTAA 230

Db 168105 GAGTATGCCGCTGTGTGGAGAGAGTGTGTGCTAGAGTCTCGGAGCCCATCGCAGCGTGA 168164
Qy 231 CCTGGATCTCCGCGAGGCTGCGGAGGCGCGCCACCTGGNGGGCATT 278
Db 168165 CCTGGATCTCCGCGGGTGTGGCGGAGGCGGGCCACCTGGAGGGGACATT 168212

Search completed: August 27, 2003, 06:41:21
Job time : 6537 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 28, 2003, 13:37:24 ; Search time 2589 seconds
(without alignments)
187.752 Million cell updates/sec

Title: US-10-042-417A-91

Perfect score: 110

Sequence: 1 NNAGSVEWTKPKGLRRROT 20

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-O=/cgn2.1/USPTO.spool/US10042417/runat_19082003_141155_27898/app_query.fasta_1.199
-DB=EST -QWTF=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTPMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10042417 -CGN_1_1_2810 -runat_19082003_141155_27898 -NCPU=3
-NO_MMAP -LARGESQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estim:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_estl:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: 3b_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	
1	91	82.7	275	9	AA361254	AA361254 EST70554	
C	2	91	82.7	343	10	BE815838	BE815838 PM3-BN017
	3	91	82.7	354	9	AA180937	AA180937 zp44f10.i.r
C	4	91	82.7	469	9	AI201163	AI201163 qf70a12.x
	5	91	82.7	619	14	CB436760	CB436760 683911.MA
6	91	82.7	783	9	AU117363	AU117363 AU117363	
7	91	82.7	842	9	AU121935	AU121935 AU121935	
8	91	82.7	918	13	BU554135	BU554135 AGENCOURT	
9	87	79.1	284	9	AA362744	AA362744 EST72502	
10	86	78.2	704	9	AV724591	AV724591 AV724591	
11	85	77.3	540	9	AV748756	AV748756 AV748756	
12	85	77.3	601	9	AU137449	AU137449 AU137449	
13	78	70.9	366	9	AA192724	AA192724 zp90h03.i.r	
14	74	67.3	439	10	BB849189	BB849189 BB849189	
15	74	67.3	530	13	BQ780318	BQ780318 UT-R-FF0-	
16	74	67.3	1884	11	AK047669	AK047669 Mus muscu	
17	74	67.3	3081	11	AK046676	AK046676 Mus muscu	
18	74	67.3	3084	11	AK050240	AK050240 Mus muscu	
19	72	65.5	421	14	CB552833	CB552833 NMSP0025	
20	72	65.5	693	14	CB552221	CB552221 NMSP0030	
21	67	60.9	508	9	AV749575	AV749575 AV749575	
22	67	60.9	920	13	BX410337	BX410337 BX410337	
23	65	59.1	220	10	BB465637	BB465637 BB465637	
24	62	56.4	889	13	BQ939384	BQ939384 AGENCOURT	
25	61	55.5	882	10	BF569872	BF569872 602185818	
C	26	60	54.5	219	28	AZ052030	AZ052030 RPCI-23-4
C	27	59	53.6	277	28	AZ071740	AZ071740 RPCI-23-4
C	28	59	53.6	507	9	AW791810	AW791810 D00819-R
C	29	58	53.6	544	28	AZ062435	AZ062435 RPCI-23-4
C	30	58	52.7	186	10	BB261934	BB261934 BB261934
C	31	58	52.7	952	13	BQ945311	BQ945311 AGENCOURT
C	32	58	52.7	1257	13	BQ674608	BQ674608 AGENCOURT
C	33	57	51.8	216	10	BB217471	BB217471 BB217471
C	34	57	51.8	282	10	BB421502	BB421502 BB421502
C	35	57	51.8	308	10	BB561032	BB561032 BB561032
C	36	57	51.8	523	14	CA498025	CA498025 WHE3237.D
C	37	57	51.8	839	10	BG489863	BG489863 602518754
C	38	57	51.8	846	10	BE737746	BE737746 601572358
C	39	57	51.8	1330	12	BM811259	BM811259 AGENCOURT
C	40	57	51.8	1374	10	BF135790	BF135790 601782077
C	41	57	51.8	1434	12	BM478638	BM478638 AGENCOURT
C	42	56	50.9	252	14	RI8972	RI8972 yg25a09.r1
C	43	56	50.9	270	10	BB529646	BB529646 BB529646
C	44	56	50.9	526	29	TA275E050	TA275E050 T. brucei
C	45	56	50.9	578	14	W85523	W85523 mf58e03.r1

ALIGNMENTS

RESULT 1
AA361254

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AA361254 275 bp mRNA linear EST 21-APR-1997
EST70554 T-cell lymphoma Homo sapiens cDNA 5' end similar to kinase
inhibitor p27kipl, cyclin-dependent, mRNA sequence.

AA361254
GI:2013615

EST

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 275)

AUTHORS

Adams, M.D., Kerlavage, A.R., Fleischmann, R.D., Fuldner, R.A., Bult, C.J., Lee, N.H., Kirkness, E.F., Weinstock, K.G., Gocayne, J.D., White, O., Sutton, G., Blake, J.A., Brandon, R.C., Man-Wai, C., Clayton, R.A., Cline, T.R., Cotton, M.D., Earle-Hughes, J., Fine, L.D., Fitzgerald, L.M., Fitzhugh, W.M., Fritchman, J.L., Geoghegan, N.S., Glodek, A., Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkle, P.S., Jr., Kelley, J.M., Kelley, J.C., Liu, L.-I., Marnaros, S.M., Merrick, J.M., Moreno-Palancas, R.F., McDonald, L.A., Nguyen, D.T., Pelligrino, S.M., Phillips, C.A., Ryder, S.E., Scott, J.L., Saudek, D.M., Shirley, R., Small, K.V., Spriggs, T.A., Utterback, T.R., Weidman, J.F., Li, Y., Bednarik, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.J., Dimke, D., Feng, D.-F., Ferrie, A., Fischer, C., Hastings, G.A., He, W.W., Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K., Kozak, D.L., Kunsch, C., Hungjun, J., Li, H., Weissner, P.S., Olsen, H., Raymond, L., Wei, Y.F., Wing, J., Xu, C., Yu, G.L., Ruben, S.M., Dillion, P.J., Fannon, M.R., Rosen, C.A., Haseltine, W.A., Fields, C., Fraser, C.M. and Venter, J.C.

TITLE

Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence

JOURNAL

MEDLINE

PUBMED

COMMENT

96026280

7566098

Other ESTs:

THC192705

Contact: Kerlavage, AR

Bioinformatics

The Institute for Genomic Research

9712 Medical Center Drive, Rockville, MD 20850 USA

Tel: 3018699056

Fax: 3018699423

Email: arkerlav@tigr.org

For clone availability, additional sequence and expression

information related to this EST, please check the TIGR Human Gene

Index (<http://www.tigr.org/tdb/hgi/hgi.html>)

Seq primer: M13 Reverse.

FEATURES

source

Location/Qualifiers

1..275

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="ATCC (inhost):165434"

/dev_stage="Adult"

/clone_lib="T-lymphocyte"

/note="Organ: T-cell lymphoma"

/vector="pBluescript SK-; Site_1: EcoRI; Site_2: XhoI"

BASE COUNT

ORIGIN

90 a 64 c 65 g 53 t 3 others

Alignment Scores:

Pred. No.: 0.00514 Length: 275
Score: 91.00 Matches: 18
Percent Similarity: 94.74% Conservatives: 0
Best Local Similarity: 94.74% Mismatches: 1
Query Match: 82.73% Indels: 0
DB: Gaps: 0

US-10-042-417a-91 (1-20) x AA361254 (1-275)

QY 2

AsnAlaclySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
|||||
156 AATGCCGGTCTCTGGAGCAGAGCCCAAGAGCCCTGGCTCAGAGACGCTCAACG 212

Db

RESULT 2

BE815838/c

LOCUS

DEFINITION

PM3-BN0174-300500-010-c02 BN0174 Homo sapiens cDNA, mRNA sequence.

BE815838

VERSION

BE815838.1 GI:10248072

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 343)

AUTHORS

Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

TITLE

Shoiguin sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

10737800

PUBMED

COMMENT

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome

Project. This entry can be seen in the following URL

(<http://www.ludwig.org.br/scripts/gethtml2.pl?ti=6t2-pm3-BN0174-300>)

500-010-c02&t3=2000-05-30&t4=1

Seq primer: puc 18 forward

High quality sequence start: 14

High quality sequence stop: 343.

Location/Qualifiers

1..343

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/dev_stage="Adult"

/clone_lib="BN0174"

/note="Organ: breast normal; Vector: puc18; Site_1: SmaI;

Site_2: SmaI; A mini-library was made by cloning products

derived from ORESTES PCR (U.S. Letters Patent application

No. 196,716 - Ludwig Institute for Cancer Research)

profiles into the pUC 18 vector. Reverse transcription of

tissue mRNA and cDNA amplification were performed under

low stringency conditions."

BASE COUNT 65 a 88 c 90 g 100 t

ORIGIN

Alignment Scores:

Pred. No.: 0.00648 Length: 343
Score: 91.00 Matches: 18
Percent Similarity: 94.74% Conservatives: 0
Best Local Similarity: 94.74% Mismatches: 1
Query Match: 82.73% Indels: 0
DB: Gaps: 0

US-10-042-417a-91 (1-20) x BE815838 (1-343)

QY 2

AsnAlaclySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
|||||
137 AATGCCGGTCTCTGGAGCAGAGCCCAAGAGCCCTGGCTCAGAGACGCTCAACG 81

Db

RESULT 3

AA180937

LOCUS

DEFINITION

2P44f10_r1 Stragatene muscle 937209 Homo sapiens cDNA clone

IMAGE:612331 5' similar to SW:CDNB_HUMAN P46527 CYCLIN-DEPENDENT

KINASE INHIBITOR 1B, mRNA sequence.

AA180937

VERSION

AA180937.1 GI:1764429

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 354)

AUTHORS

Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin

J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wylie, T., Waterston, R., and Wilson, R.
 WashU-NCI Human EST Project
 Unpublished
 Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu

This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 520 Std Error: 0.00
 Seq primer: -28M13 rev2 from AmerSham
 High quality sequence stop: 288.

FEATURES

Location/Qualifiers
 1..354
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="GDB:4644170"
 /db_xref="taxon:9606"
 /clone="IMAGE:612331"
 /tissue_type="muscle"
 /dev_stage="adult"
 /lab_host="SOLR (kanamycin resistant)"
 /clone_lib="Stratagene muscle 937209"
 /note="Organ: skeletal muscle; Vector: pBluescript SK-;
 Site 1: EcoRI; Site 2: XhoI; Cloned unidirectionally.
 Primer: Oligo dT. Skeletal muscle from patient with
 malignant hyperthermia. Average insert size: 1.0 kb;
 Uni-ZAP XR Vector; ~5' adaptor sequence: 5' GAATTCGGCAGG
 3' -3' adaptor sequence: 5' CTCAGTGTGTGTGTGTGTGTGTGT 3'

BASE COUNT 118 a 69 c 60 g 107 t
 ORIGIN

Alignment Scores:

Pred. No.: 0.0067 Length: 354
 Score: 91.00 Matches: 18
 Percent Similarity: 94.74% Conservative: 0
 Best Local Similarity: 94.74% Mismatches: 1
 Query Match: 82.73% Indels: 0
 DB: 9 Gaps: 0

US-10-042-417a-91 (1-20) x AAL80937 (1-354)

QY 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
 Db 8 AATGCCGGTCTCTGTGGAGCAGCGCCCAAGAGCGCTGGCTCAGAACGCTCAACG 64

RESULT 4

AI201163/c
 LOCUS 469 bp mRNA linear EST 14-OCT-1998
 DEFINITION gf70a12.x1 Soares testis_NHT Homo sapiens cDNA clone IMAGE:1755358
 3' similar to SW:CDNE_HUMAN P46527 CYCLIN-DEPENDENT KINASE

INHIBITOR 1B 1, mRNA sequence.

AI201163

AI201163.1 GI:3753769

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 469)

NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished

Contact: Robert Strausberg, Ph.D.

Email: cgapb-r@mail.nih.gov

cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo

, Ph.D.

cDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CCGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/bbrp/image/image.html
 Seq primer: -40UP from Gibco
 High quality sequence stop: 426.

FEATURES

Location/Qualifiers
 1..469
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:1755358"
 /sex="male"
 /lab_host="DH10B"

/clone_lib="Soares testis_NHT"

/note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
 was prepared from mRNA obtained from Clontech Laboratories
 , Inc., and primed with a Not I - oligo(dT) primer [5'
 TGTACCAATCTCAAGTGGGAGCGGCCCAATTTTTTTTTTTT 3'].
 Double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT73 vector. Library
 went through one round of normalization to Cot5, and was
 constructed by Bento Soares and M. Fatima Bonaldo."

-BASE COUNT 107 a 89 c 96 g 177 t

ORIGIN

Alignment Scores:
 Pred. No.: 0.009 Length: 469
 Score: 91.00 Matches: 18
 Percent Similarity: 94.74% Conservative: 0
 Best Local Similarity: 94.74% Mismatches: 1
 Query Match: 82.73% Indels: 0
 DB: 9 Gaps: 0

US-10-042-417a-91 (1-20) x AI201163 (1-469)

QY 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20

Db 253 AATGCCGGTCTCTGTGGAGCAGCGCCCAAGAGCGCTGGCTCAGAACGCTCAACG 197

RESULT 5

CB436760

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Bos taurus

Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Iutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

Bovidae; Bovinae; Bos.

1 (bases 1 to 619)

Smith, T.P.I., Roberts, A.J., Ehternkamp, S.E., Chitko-McKown, C.G.,

Wray, J.E., and Keele, J.W.

A second set of bovine ESTs from pooled-tissue normalized libraries

Unpublished

Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center

PO Box 166, Clay Center, NE 68933-0166, USA

Tel: 402 762 4366

Fax: 402 762 4390

Email: smith@mail.marc.usda.gov

Single pass sequencing. Bases called with phred v0.020425.c and

trimmed with the aid of the trim_ait option. Vector identified with

cross_match v0.990329.

Plate: FOYE038 row: J column: 8

Seq primer: GTAATACGACTCATATAGG.

Location/Qualifiers

1..619

/organism="Bos taurus"

/mol_type="mRNA"

source

BASE COUNT 194 a 114 c 135 g 174 t 2 others
 ORIGIN

Alignment Scores:
 Pred. No.: 0.012 Length: 619
 Score: 91.00 Matches: 18
 Percent Similarity: 94.74% Conservative: 0
 Best Local Similarity: 94.74% Mismatches: 1
 Query Match: 82.73% Indels: 0
 DB: 14 Gaps: 0

US-10-042-417a-91 (1-20) x CB436760 (1-619)

QY 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
 Db 79 AACGCCGGTCTGTGGAGCAGACGCCCAAGAGCGCTGGCTCAGAGCGTCAGACG 135

RESULT 6

LOCUS AUI17363 HEMBA1 Homo sapiens cDNA clone HEMBA1001231 5', mRNA EST 01-AUG-2002
 DEFINITION

ACCESSION AUI17363
 VERSION AUI17363.1 GI:10932325
 KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 783)
 Ota.T., Nishikawa.T., Suzuki.Y., Ishii.S., Saito.K., Kawai.Y.,
 Yamamoto.J., Wakamatsu.A., Nakamura.Y., Nagai.T., Sugano.S. and
 Isogai.T.

HRI human cDNA project

TITLE

JOURNAL

COMMENT

Contact: Takao Isogai
 Genomics Laboratory
 Helix Research Institute
 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
 Tel: 81-438-52-3975
 Fax: 81-438-52-3986

Email: genomics@hri.co.jp
 HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
 Research Institute; cDNA library construction: Department of
 Virology, Institute of Medical Science, University of Tokyo, and
 Helix Research Institute.

FEATURES

source

1..783
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="HEMBA1001231"
 /tissue_type="whole embryo, mainly head"
 /dev_stage="embryo, 10 weeks"
 /clone_lib="HEMBA1"
 /note="Vector: pME18SFL3"
 BASE COUNT 237 a 177 c 196 g 170 t 3 others
 ORIGIN

Alignment Scores:

Pred. No.: 0.0154 Length: 783
 Score: 91.00 Matches: 18
 Percent Similarity: 94.74% Conservative: 0
 Best Local Similarity: 94.74% Mismatches: 1
 Query Match: 82.73% Indels: 0

DB: 9 Gaps: 0
 US-10-042-417a-91 (1-20) x AUI17363 (1-783)

QY 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
 Db 419 AATGCCGGTCTGTGGAGCAGACGCCCAAGAGCGCTGGCTCAGAGAGCGTCAACG 475

RESULT 7

LOCUS AUI21935 MAMMA1 Homo sapiens cDNA clone MAMMA1001309 5', mRNA EST 01-AUG-2002
 DEFINITION

ACCESSION AUI21935
 VERSION AUI21935.1 GI:10937170
 KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 842)
 Ota.T., Nishikawa.T., Suzuki.Y., Ishii.S., Saito.K., Kawai.Y.,
 Yamamoto.J., Wakamatsu.A., Nakamura.Y., Nagai.T., Sugano.S. and
 Isogai.T.

HRI human cDNA project

TITLE

JOURNAL

COMMENT

Contact: Takao Isogai
 Genomics Laboratory
 Helix Research Institute
 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
 Tel: 81-438-52-3975
 Fax: 81-438-52-3986

Email: genomics@hri.co.jp
 HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
 Research Institute; cDNA library construction: Department of
 Virology, Institute of Medical Science, University of Tokyo, and
 Helix Research Institute.

FEATURES

source

1..842
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MAMMA1001309"
 /tissue_type="mammary gland"
 /clone_lib="MAMMA1"
 /note="Vector: pME18SFL3"
 BASE COUNT 245 a 202 c 234 g 159 t 2 others
 ORIGIN

Alignment Scores:

Pred. No.: 0.0166 Length: 842
 Score: 91.00 Matches: 18
 Percent Similarity: 94.74% Conservative: 0
 Best Local Similarity: 94.74% Mismatches: 1
 Query Match: 82.73% Indels: 0
 DB: 9 Gaps: 0

US-10-042-417a-91 (1-20) x AUI21935 (1-842)

QY 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
 Db 584 AATGCCGGTCTGTGGAGCAGACGCCCAAGAGCGCTGGCTCAGAGAGCGTCAACG 640

RESULT 8

LOCUS BU554135 AGENCOURT_10368772 NIH_MGC_109 Homo sapiens cDNA clone IMAGE:6581386 5', mRNA sequence.
 DEFINITION

ACCESSION BU554135
 VERSION BU554135.1 GI:22904407
 KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

source

Location/Qualifiers

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="BU554135"

/tissue_type="whole embryo, mainly head"

/dev_stage="embryo, 10 weeks"

/clone_lib="HEMBA1"

/note="Vector: pME18SFL3"

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 918)

NIH-MGC http://mgc.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-rt@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory

DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: L1CM2787 row: n column: 10

High quality sequence stop: 575.

Location/Qualifiers

1. .918

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6581386"

/tissue_type="teratocarcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_109"

/note="Organ: ovary; Vector: pOTB7; Site:1; ECORI; Site:2;

XhoI; cDNA made by oligo-dT priming. Directionally cloned

into ECORI/XhoI sites using the following 5' adaptor:

GGCAGGAG(G). Library constructed by Ling Hong in the

Laboratory of Gerald M. Rubin (University of California,

Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and

Superscript II RT (Life Technologies). Note: this is a

NIH_MGC Library."

BASE COUNT 242 a 234 c 268 g 173 t 1 others

ORIGIN

Alignment Scores:

Pred. No.: 0.0182 Length: 918

Score: 91.00 Matches: 18

Percent Similarity: 94.74% Conservativeness: 0

Best Local Similarity: 94.74% Mismatches: 1

Query Match: 82.73% Indels: 0

DB: 13 Gaps: 0

US-10-042-417A-91 (1-20) x BU554135 (1-918)

QY 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20

Db 649 AATGCCGGTCTGTGGAGCAGACGCCCAAGAGCGCTGCCTCAGACGCTCAACG 705

RESULT 9

AA362744

LOCUS

EST72502 Ovary II Homo sapiens cDNA 5' end similar to kinase

inhibitor p27kip1, cyclin-dependent, mRNA sequence.

AA362744

VERSION

AA362744.1 GI:2015064

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 284)

Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult

,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White

,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C., Clayton,R.A.,

Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald

,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S., Glodek,A.,

Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr., Kelley,J.M.,

Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,

Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,

Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,

Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
Bednarik,P.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A., He,W.W.,
Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kozak,D.L.,
Kunsch,C., Wang,J., Li,H., Weissner,P.S., Olsen,H., Raymond,L.,
Wei,Y.F., Hing,J., Xu,C., Yu,G.L., Ruben,S.M., Dillion,P.J., Fannon
M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and
Venter,J.C.

Initial assessment of human gene diversity and expression patterns
based upon 83 million nucleotides of cDNA sequence

Nature 377 (6547 Suppl), 3-174 (1995)

96026280

PUBLISHED

7566098

Other ESTs: THC192705

Contact: Kerlavage, AR

Bioinformatics

The Institute for Genomic Research

9712 Medical Center Drive, Rockville, MD 20850 USA

Tel: 3018699056

Fax: 3018699423

Email: arkerlavet@tigr.org

For clone availability, additional sequence and expression

information related to this EST, please check the TIGR Human Gene

Index (<http://www.tigr.org/tdb/hgi/hgi.html>)

Seq primer: M13 Reverse.

FEATURES

Location/Qualifiers

1. .284

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="ATCC (inhost):167028"

/db_xref="taxon:9606"

/sex="female"

/dev_stage="adult"

/clone_lib="Ovary II"

/note="Organ: ovary; Vector: pBluescript SK-; Site:1;

ECORI; Site:2: XhoI"

BASE COUNT 98 a 59 c 60 t 7 others

ORIGIN

Alignment Scores:

Pred. No.: 0.0186 Length: 284

Score: 87.00 Matches: 17

Percent Similarity: 89.47% Conservativeness: 0

Best Local Similarity: 89.47% Mismatches: 2

Query Match: 79.09% Indels: 0

DB: 9 Gaps: 0

US-10-042-417A-91 (1-20) x AA362744 (1-284)

QY 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20

Db 69 AATNCCGGTCTGTGGAGCAGACGCCCAAGAGCGCTGCCTCAGACGCTCAACG 155

RESULT 10

AV724591

LOCUS

DEFINITION

AV724591 HTB Homo sapiens cDNA clone HTBAG06 5', mRNA sequence.

AV724591

ACCESSION

AV724591.1 GI:10829130

VERSION

EST.

KEYWORDS

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 704)

Gu,Y., Peng,Y., Song,H., Huang,Q., Yang,Y., Gao,G., Xiao,H., Xu,X.,

Li,N., Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng,L., Xu

,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Hu,R.,

Chen,J., Chen,Z. and Han,Z.

Homo sapiens cDNA HTB clones

Unpublished

CONTACT: Ziguang Han

Chinese National Human Genome Center at Shanghai

351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES

source
1. .704

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HTBAG06"
/tissue_type="Hypothalamus"
/dev_stage="Adult"
/lab_host="SOLR"
/clone_lib="HTB"

/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT 217 a 115 c 151 g 221 t
ORIGIN

Alignment Scores:

Pred. No.: 0.0661 Length: 704
Score: 86.00 Matches: 17
Percent Similarity: 94.44% Conservative: 0
Best Local Similarity: 94.44% Mismatches: 1
Query Match: 78.18% Indels: 0
DB: Gaps: 0

US-10-042-417A-91 (1-20) x AV724591 (1-704)

QY 2 AsnAlaGlySerValGluTrpThrProLysProGlyLeuArgArgGln 19

|||||
Db 31 AATGCCGGTCTCTGGAGCAGACGCCCAAGAGCCTGGCCTCAGAGACGTCAA 84

RESULT 11

AV748756
LOCUS AV748756 NPC Homo sapiens cDNA clone NPCA0H04 5', mRNA linear EST 19-OCT-2000
DEFINITION AV748756
ACCESSION AV748756
VERSION AV748756.1 GI:10906604

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS Song,H., Peng,Y., Gu,Y., Yang,Y., Gao,G., Xiao,H., Xu,X., Li,N.,
Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng,L., Xu,S., Gu
,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Ye,M., Zhang,Q.
, Han,Z., Chen,Z., Hu,R. and Chen,J.

Homo sapiens NPC library cDNA clones

Unpublished

Contact: Qinghua Zhang

Shanghai Institute of Endocrinology, Rui-Jin Hospital

197 Rui-Jin II Road, Shanghai 200025, P. R. China

Tel: 86-21-64370045(ex.663332)

Fax: 86-21-64743206

Email: mbshi@ms.sh.cn

This clone is available at Shanghai Hematology Institute in

Shanghai.

Chinese National Human Genome Center at Shanghai

351 Guo Shoujing Road, Zhangjiang Hi- Tech Park, Pudong.

FEATURES

source
1. .540

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NPCA0H04"
/tissue_type="pituitary"
/dev_stage="Adult"
/lab_host="SOLR"
/clone_lib="NPC"

/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT 164 a 119 c 131 g 121 t 5 others
ORIGIN

Alignment Scores:

Pred. No.: 0.0684 Length: 540
Score: 85.00 Matches: 17
Percent Similarity: 89.47% Conservative: 0
Best Local Similarity: 89.47% Mismatches: 2
Query Match: 77.27% Indels: 0
DB: Gaps: 0

US-10-042-417A-91 (1-20) x AV748756 (1-540)

QY 2 AsnAlaGlySerValGluTrpThrProLysProGlyLeuArgArgGlnThr 20

|||||
Db 245 AATGCCGGTCTCTGGAGCAGACGCCCAAGAGCCTGGCCTCAGAGACGTCAAACG 301

RESULT 12

AU137449
LOCUS AU137449 PLACE1 Homo sapiens cDNA clone PLACE1006501 5', mRNA
DEFINITION AU137449
ACCESSION AU137449
VERSION AU137449.1 GI:10997988

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
Isogai,Y.

HRI human cDNA project

Unpublished

JOURNAL

COMMENT Contact: Takao Isogai

Genomics Laboratory

Helix Research Institute

1532-3 Yana, Kisarazu, Chiba 292-0812, Japan

Tel: 81-438-52-3975

Fax: 81-438-52-3986

Email: genomics@hri.co.jp

HRI human cDNA project; 5'- & 3'-end one pass sequencing; Helix
Research Institute; cDNA library construction; Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.

FEATURES

source
1. .601

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="PLACE1006501"

/tissue_type="placenta"

/clone_lib="PLACE1"

/note="Vector: pME18SFL3"

BASE COUNT 166 a 166 c 179 g 87 t 3 others
ORIGIN

Alignment Scores:

Pred. No.: 0.0766 Length: 601
Score: 85.00 Matches: 17
Percent Similarity: 89.47% Conservative: 0
Best Local Similarity: 89.47% Mismatches: 2
Query Match: 77.27% Indels: 0
DB: Gaps: 0

US-10-042-417A-91 (1-20) x AU137449 (1-601)

QY 2 AsnAlaGlySerValGluTrpThrProLysProGlyLeuArgArgGlnThr 20

|||||
Db 532 AATGCCGGTCTCTGGAGCAACGCCCAAAACCTGGCCTCANAAGACGTCAAACG 588


```

RESULT 13
AA192724
LOCUS
DEFINITION
    zp0h03.r1 Stratagene HeLa cell s3 937216 Homo sapiens cDNA clone
    IMAGE:627509 5' similar to SW:CDNB_HUMAN P46527 CYCLIN-DEPENDENT
    KINASE INHIBITOR 1B ; mRNA sequence.
ACCESSION
VERSION
    AA192724
SOURCE
    Homo sapiens (human)
ORGANISM
REFERENCE
    1 (bases 1 to 366)
AUTHORS
    Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B.,
    Chissole, S., Dietrich, N., DuBuque, T., Favell, A., Gish, W., Hawkins
    M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N., Mardis, E., Moore
    B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T.,
    Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J., Trevisakis, E.,
    Underwood, K., Wohlmann, P., Waterston, R., Wilson, R. and Marra, M.
    Generation and analysis of 280,000 human expressed sequence tags
    Genome Res. 6 (9), 807-828 (1996)
TITLE
    97044478
JOURNAL
MEDLINE
PUBMED
    8889549
COMMENT
    Contact: Wilton RK
    Washington University School of Medicine
    4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
    Tel: 314 286 1800
    Fax: 314 286 1810
    Email: est@wustl.edu
    This clone is available royalty-free through LLNL ; contact the
    IMAGE Consortium (info@image.llnl.gov) for further information.
    Seq primer: -28M13 rev2 from Amersham
    High quality sequence stop: 225.
FEATURES
    source
        1..366
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="GDB:5046207"
        /db_xref="taxon:9606"
        /clone="IMAGE:627509"
        /sex="female"
        /dev_stage="HeLa S3 cell line"
        /lab_host="SOLR (kanamycin resistant)"
        /clone_lib="Stratagene HeLa cell s3 937216"
        /note="Vector: pBluescript SK-; Site.1: EcoRI; Site.2:
        XhoI; Cloned unidirectionally. Primer: Oligo dt. HeLa S3
        epithelioid carcinoma cells grown to semi-confluency
        without induction. Average insert size: 1.5 kb; Uni-ZAP XR
        vector. -5' adaptor sequence: 5' GAATTCGGCAGCAG 3' -3'
        adaptor sequence: 5' CTCGAGTGTGTGTGTGTGTGTGT 3'"
        134 a 81 c 76 g 74 t 1 others
BASE COUNT
ORIGIN
    Alignment Scores:
    Pred. No.: 0.408 Length: 366
    Score: 78.00 Matches: 18
    Percent Similarity: 90.00% Conservative: 0
    Best Local Similarity: 90.00% Mismatches: 1
    Query Match: 70.91% Indels: 1
    DB: 9 Gaps: 0
US-10-042-417A-91 (1-20) x AA192724 (1-366)
QY 2 AsnAlaGlySerValGluTrpThrProLysLys-ProGlyLeuArgArgGlnThr 20
    |||||
Db 139 AATGCCGGTCTGTGTGAGCAGAGCCCAAGAGCCCTGGCCCTCAGAAGAGCGTCAACG 196
    |||||
RESULT 14
BB849189
LOCUS
DEFINITION
    BB849189 439 bp mRNA linear EST 26-NOV-2001
    BB849189 RIKEN full-length enriched, adult inner ear Mus musculus

```

```

ACCESSION
VERSION
    BB849189.1
KEYWORDS
    EST.
SOURCE
    Mus musculus (house mouse)
ORGANISM
REFERENCE
    1 (bases 1 to 439)
AUTHORS
    Akimura, T., Arakawa, T., Carninci, P., Furuno, M., Hanagaki, T.,
    Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Imotani, K., Ishii
    J., Ito, M., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Matsuyama, T.,
    Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Okazaki, Y., Okido, T.,
    Saito, R., Sakai, C., Sakai, K., Sakazume, N., Sasaki, D., Sato, K.,
    Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagawa
    A., Takatahshi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T.,
    Watahiki, A., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
    RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al.
    2001)
JOURNAL
COMMENT
    Contact: Yoshihide Hayashizaki
    Laboratory for Genome Exploration Research Group, RIKEN Genomic
    Sciences Center(GSC), Yokohama Institute
    The Institute of Physical and Chemical Research (RIKEN)
    1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
    Tel: 81-45-503-9222
    Fax: 81-45-503-9216
    Email: genome-res@gsc.riken.go.jp,
    URL:http://genome.gsc.riken.go.jp/
    Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh
    M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
    Normalization and subtraction of cap-trapper-selected cDNAs to
    prepare full-length cDNA libraries for rapid discovery of new
    genes. Genome Res. 10 (10), 1617-1630 (2000)
    wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
    Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura
    S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and
    Hayashizaki, Y.
    RIKEN integrated sequence analysis (RISA) system--384-format
    sequencing pipeline with 384 multicapillary sequencer. Genome Res.
    10 (11), 1757-1771 (2000)
    Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara
    Y. and Hayashizaki, Y.
    Computer-based methods for the mouse full-length cDNA
    encyclopedia: real-time sequence clustering for construction of a
    nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
    Please visit our web site (http://genome.gsc.riken.go.jp) for
    further details.
    e mouse tissues.
FEATURES
    source
        1..439
        /organism="Mus musculus"
        /mol_type="mRNA"
        /strain="C57BL/6J"
        /db_xref="taxon:10090"
        /clone="F930006P16"
        /tissue_type="inner ear"
        /dev_stage="adult"
        /clone_lib="RIKEN full-length enriched, adult inner ear"
        144 a 96 c 96 g 103 t
BASE COUNT
ORIGIN
    Alignment Scores:
    Pred. No.: 1.73 Length: 439
    Score: 74.00 Matches: 14
    Percent Similarity: 94.12% Conservative: 2
    Best Local Similarity: 82.35% Mismatches: 1
    Query Match: 67.27% Indels: 0
    DB: 10 Gaps: 0
US-10-042-417A-91 (1-20) x BB849189 (1-439)
QY 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgArg 18

```

cDNA clone: F930006P16 5', mRNA sequence.

```

Db      149 AACGCTGGCACTGTGGAGCAGAGCCCAAGAGCGCGCTTCGACGCCAG 199
|||||
RESULT 15
BQ780318/c
LOCUS
DEFINITION
  BQ780318 530 bp mRNA linear EST 26-JUL-2002
  UI-R-FF0-cpb-p-13-0-UI.s1 UI-R-FF0 Rattus norvegicus cDNA clone
  UI-R-FF0-cpb-p-13-0-UI 3', mRNA sequence.
ACCESSION
  BQ780318
VERSION
  BQ780318.1 GI:21988790
KEYWORDS
  EST.
SOURCE
  Rattus norvegicus (Norway rat)
  Rattus norvegicus
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
  Rattus.
REFERENCE
  1 (bases 1 to 530)
  Bonaldo,M.F., Lennon,G. and Soares,M.B.
  Normalization and subtraction: two approaches to facilitate gene
  discovery
  Genome Res. 6 (9), 791-806 (1996)
  9704477
  889548
  Contact: Soares, MB
  Coordinated Laboratory for Computational Genomics
  University of Iowa
  375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
  Tel: 319 335 8250
  Fax: 319 335 9565
  Email: bento-soares@iowa.edu
  Tissue Procurement: Jeff Stevens
  cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
  cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
  DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
  Clone Distribution: DISTRIBUTION: Researchers may obtain clones
  from Research Genetics (www.resgen.com).
  Seq primer: M13 FORWARD
  POLYA=Yes.

```

FEATURES

```

source
  1..530
  Location/Qualifiers
    /organism="Rattus norvegicus"
    /mol_type="mRNA"
    /db_xref="taxon:10116"
    /clone="UI-R-FF0-cpb-p-13-0-UI"
    /tissue_type="Mixed tissues"
    /dev_stage="Adult"
    /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
    /clone_lib="UI-R-FF0"
    /note="Vector: pT7T3-Pac (Pharmacia) with a modified
    polylinker; Site 1: EcoR I; Site 2: Not I; UI-R-FF0 is a
    subtracted cDNA library containing the following tissue(s)
    ): Normal cartilage and SR-JWS Tumor Line. The
    subtraction was made according to Bonaldo, Lennon and
    Soares, Genome Research, 6:791-806, 1996. The
    oligonucleotide used to prime the synthesis of
    first-strand cDNA contains a library tag sequence that is
    located between the Not I site and the (dT)18 tail. The
    sequence tags for these libraries are: CTAATGGAGC,
    CATTCTTGTA.
  TAG_LIB=UI-R-FF0
  TAG_TISSUE=rat SRC-JWST tumor line
  TAG_SEQ=CATTCTTGTA

```

```

BASE COUNT 98 a 143 c 129 g 159 t 1 others
ORIGIN

```

Alignment Scores:

```

Pred. No.: 2.11 Length: 530
Score: 74.00 Matches: 14
Percent Similarity: 94.12% Conservative: 2
Best Local Similarity: 82.35% Mismatches: 1
Query Match: 67.27% Indels: 0
DB: 13 Gaps: 0

```

```

US-10-042-417A-91 (1-20) x BQ780318 (1-530)

```

```

QY      2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgArg 18
|||||
DB      213 AATGCTGGCACTGTGGAGCAGAGCCCAAGAGCGCGCTTCGACGCCAG 163
|||||

```

```

Search completed: August 28, 2003, 15:17:14
Job time : 2597 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2003, 09:19:44 ; Search time 204 Seconds
(without alignments)

25.299 Million cell updates/sec

Title: US-10-042-417A-91

Perfect score: 110
Sequence: 1 NNAGSVETPKPKGLRRQT 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL_23.*

1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mmc.*
8: sp_organelle.*
9: sp_phase.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	91	82.7	198	4 Q9NYG6	Q9NYG6 homo sapien
2	91	82.7	198	4 Q96TE0	Q96TE0 homo sapien
3	82	74.5	198	6 Q9BDC3	Q9BDC3 sus scrofa
4	74	67.3	197	11 Q35792	Q35792 rattus norv
5	74	67.3	197	11 Q08769	Q08769 rattus norv
6	74	67.3	197	11 Q8BG74	Q8BG74 mus musculus
7	55	50.0	134	16 Q8ET29	Q8ET29 shewanela
8	54	49.1	198	13 Q8JIV2	Q8JIV2 gallus gall
9	54	49.1	874	5 Q18109	Q18109 caenorhabdi
10	53	48.2	509	3 Q9P4V7	Q9P4V7 candida boi
11	48	43.6	418	2 Q8VTA4	Q8VTA4 agrobacteri
12	48	43.6	450	2 Q9F5A4	Q9F5A4 agrobacteri
13	48	43.6	735	16 Q9A9V4	Q9A9V4 caulobacter
14	47	42.7	391	2 Q9R6D1	Q9R6D1 agrobacteri
15	47	42.7	419	16 Q50242	Q50242 agrobacteri
16	46	41.8	178	8 Q8MLE8	Q8MLE8 chaetosphae

ALIGNMENTS

RESULT 1

Q9NYG6 PRELIMINARY; PRT; 198 AA.
ID Q9NYG6
AC Q9NYG6;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Cyclin-dependent kinase inhibitor p27kip1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Hepatoma;
RA Li J., Wang W.-L., Yang X.-K., Yu X.-X.;
RT "Homo sapiens cyclin-dependent kinase inhibitor p27kip1 cDNA."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF247551; AAF69497.1;
DR InterPro; IPR003175; CDI.
DR Pfam; PF02234; CDI; 1.
KW Kinase.

SQ SEQUENCE 198 AA: 22099 MW; 40F7D58901CDF3EA CRC64;

Query Match 82.7%; Score 91; DB 4; Length 198;
Best Local Similarity 94.7%; Pred. No. 2.2e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 NAGSVETPKPKGLRRQT 20

DB 180 NAGSVETPKPKGLRRQT 198

||||| |

||||| |

||||| |

||||| |

||||| |

||||| |

||||| |

||||| |

||||| |

||||| |

||||| |

||||| |

||||| |

||||| |

||||| |

||||| |

```

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OX Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Spleen;
RX MEDLINE=97361761; PubMed=9218722;
RA Nomura H., Sawada Y., Fujinaga K., Ohtaki S.;
RT "Cloning and characterization of rat p27Kip1, a cyclin-dependent
RP kinase inhibitor.";
RL Gene 191:211-218(1997).
DR EMBL; D83792; BA021561.1;
DR InterPro; IPR003175; CDI.
DR Pfam; PF02234; CDI; 1.
SQ SEQUENCE 197 AA; 21112 MW; 55738078C2D55B2 CRC64;

Query Match 67.3%; Score 74; DB 11; Length 197;
Best Local Similarity 82.4%; Pred. No. 0.001;
Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 NAGSVETPKKPGGLRRR 18
    |||:| | |||||
Db 180 NAGTVEQTPKKPGLRRQ 196

RESULT 5
O08769 PRELIMINARY; PRT; 197 AA.
AC O08769;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE P27 Kip1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OX Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Kawada M., Yamagoe S., Uehara Y.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Kawada M., Yamagoe S., Kazuo S., Mizuno S., Uehara Y.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Kawada M., Yamagoe S., Murakami Y., Suzuki K., Mizuno S., Uehara Y.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY;
RA Dastvan F., Reidy M.A.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; D86924; BA019960.1;
DR EMBL; AF015194; AA071368.1;
DR InterPro; IPR003175; CDI.
DR Pfam; PF02234; CDI; 1.
SQ SEQUENCE 197 AA; 21139 MW; 55738078C2C9847F CRC64;

Query Match 67.3%; Score 74; DB 11; Length 197;
Best Local Similarity 82.4%; Pred. No. 0.001;
Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 NAGSVETPKKPGGLRRR 18
    |||:| | |||||
Db 180 NAGTVEQTPKKPGLRRQ 196

RESULT 6
Q08G74 PRELIMINARY; PRT; 197 AA.
ID Q08G74
AC Q08G74;

```

```

DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DE Cyclin-dependent kinase inhibitor 1B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Adipose tissue, Liver, and Striatum;
RX MEDLINE=22354683; PubMed=12466851;
RA The PANTOM Consortium.
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR ENBL; AK046676; BAC32833.1; -.
DR ENBL; AK047669; BAC33119.1; -.
DR ENBL; AK050240; BAC34141.1; -.
SQ SEQUENCE 197 AA; 22193 MW; BAC30D648B9BA3D6 CRC64;

Query Match 67.3%; Score 74; DB 11; Length 197;
Best Local Similarity 82.4%; Pred. No. 0.001;
Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 NAGSVEWTPKPKGLRRR 18
Db 180 NAGTVEQTPKPKGLRRQ 196

RESULT 7
Q8EI29 PRELIMINARY; PRT; 134 AA.
AC Q8EI29;
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DE NADH dehydrogenase I, A subunit.
GN NUOA OR S01021.
OS Shewanella oneidensis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Alteromonadaceae; Shewanella.
OX NCBI_TaxID=70863;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MR-1;
RX MEDLINE=22297686; PubMed=12368813;
RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
RA Read T.D., Eisen J.A., Seshadri R., Ward N., Methe B., Clayton R.A.,
RA Meyer T., Tsapin A., Scott J., Beanan M., Brinkac L., Daugherty S.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,
RA Madupu R., Peterson J.D., Umayam L.A., White O., Wolf A.M.,
RA Vamathevan J., Weidman J., Impraim M., Lee K., Berry K., Lee C.,
RA Mueller J., Khouri H., Gill J., Utterback T.R., McDonald L.A.,
RA Feldblyum T.V., Smith H.O., Venter J.C., Neilson K.H., Fraser C.M.;
RT "Genome sequence of the dissimilatory metal ion-reducing bacterium
RT Shewanella oneidensis."
RL Nat. Biotechnol. 20:1118-1123(2002).
DR ENBL; AE015546; AAN54094.1; -.
DR TIGR; S01021; -.
SQ Complete proteome.
KW SEQUENCE 134 AA; 14897 MW; 52666C2A2D12ADFA CRC64;

Query Match 50.0%; Score 55; DB 16; Length 134;
Best Local Similarity 47.1%; Pred. No. 0.67;
Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy 4 GSVEWTPKPKGLRRRQT 20
Db 117 GALEWSPKPKLNKNT 133

```

```

RESULT 8
Q8JIV2 PRELIMINARY; PRT; 198 AA.
AC Q8JIV2;
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DE Cyclin-dependent kinase inhibitor p27 Kipl.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Chen D., Boulme F., Dolznig H., Mullner E.W.;
RT "Translational regulation of p27 Kipl during terminal differentiation
RT of chicken and murine erythroid cells."
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR ENBL; AY094490; AAM22491.1; -.
DR InterPro; IPR003175; CDI.
DR Pfam; PF02234; CTI; 1.
KW Kinase.
SQ SEQUENCE 198 AA; 22260 MW; 4C650EC9EF7B6992 CRC64;

Query Match 49.1%; Score 54; DB 13; Length 198;
Best Local Similarity 63.2%; Pred. No. 1.5;
Matches 12; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 2 NAGSVEWTPKPKGLRRRQT 20
Db 180 SASSVEQTPKPKSPRRHQT 198

RESULT 9
Q18109 PRELIMINARY; PRT; 874 AA.
ID Q18109;
AC Q18109;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Hypothetical 97.0 kDa protein.
GN C23F12.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium."
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2
RA Fulton B.;
RT "The sequence of C. elegans cosmid C23F12."
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A
RC STRAIN=Bristol N2;
RA Waterston R.;
RT "Direct Submission."
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR ENBL; U39676; AAB37032.1; -.
DR HSP; P46939; IQAC.
DR WormPep; C23F12.2; CE04046.
DR InterPro; IPR001589; Actbind_actnin.
DR InterPro; IPR001715; Calponin-like.
DR InterPro; IPR001298; Filamin.

```

DR InterPro: IPR003006; Ig_MHC.
 DR Pfam: PF00307; CH; 3.
 DR Pfam: PF00630; Filamin; 3.
 DR SMART: SM00033; CH; 3.
 DR SMART: SM00557; IG_FLMN; 2.
 DR PROSITE: PS00019; ACTININ_1; 1.
 DR PROSITE: PS00021; CH; 2.
 DR PROSITE: PS01194; FILAMIN_REPEAT; 1.
 DR PROSITE: PS00290; IG_MHC; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 874 AA; 96987 MW; 2E31B2CEFFAF3BD CRC64;

Query Match 49.1%; Score 54; DB 5; Length 874;
 Best Local Similarity 80.0%; Pred. No. 7;
 Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 6 VENTPKPKGL 15
 :|||||:
 Db 815 IEWTPKEPGL 824

RESULT 10

Q9P4V7 ID Q9P4V7 PRELIMINARY; PRT; 509 AA.
 AC Q9P4V7
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Acetylsermidine oxidase.
 OS Candida boidinii (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
 OX NCBI_TaxID=5477;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20374938; PubMed=10913603;
 RA Nishikawa M., Hagishita T., Yurimoto H., Kato N., Sakai Y.,
 RA Hatanaka T.;
 RT "Primary structure and expression of peroxisomal acetylsermidine
 RT oxidase in the methylotrophic yeast Candida boidinii.";
 RL FEBS Lett. 476:150-154(2000).
 DR EMBL: AB018223; BAA97653.1; -;
 DR InterPro: IPR002937; Amino_oxidase.
 DR Pfam: PF01593; Amino_oxidase; 1.
 SQ SEQUENCE 509 AA; 57168 MW; D1EG3B699BD69EDE CRC64;

Query Match 48.2%; Score 53; DB 3; Length 509;
 Best Local Similarity 61.5%; Pred. No. 5.7;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 NNAGSVEWTPKPK 13
 :||:|:
 Db 274 NEVGSIEWSPKLP 286

RESULT 11

Q8VTA4 ID Q8VTA4 PRELIMINARY; PRT; 418 AA.
 AC Q8VTA4;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Virh.
 GN VIRH.
 OS Agrobacterium tumefaciens.
 OC Plasmid pTIAB2/73.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Rhizobiaceae; Rhizobium.
 OX NCBI_TaxID=358;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AB2/73;
 RA Schmidt J., Hammann P., Otten L.;

RT "Comparison of the vir regions of the Agrobacterium tumefaciens
 RT limited host range strain AB2/73 with those of broad host range
 RT Agrobacterium strains.";
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 DR EMBL: AF329849; AAL57007.1; -;
 DR InterPro: IPR001128; Cytochrome_P450.
 DR Pfam: PF00067; P450; 1.
 DR PRINTS: PR00385; P450.
 DR PROSITE: PS00086; CYTOCHROME_P450; 1.
 KW Heme; Monooxygenase; Oxidoreductase; Plasmid.
 SQ SEQUENCE 418 AA; 47005 MW; D75547B71D69F51F CRC64;

Query Match 43.6%; Score 48; DB 2; Length 418;
 Best Local Similarity 66.7%; Pred. No. 28;
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 7 EMTPKPKGLRRR 18
 :||:|:
 Db 41 ENRPKRPFLLRQ 52

RESULT 12

Q9F5A4 ID Q9F5A4 PRELIMINARY; PRT; 450 AA.
 AC Q9F5A4;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Riorf151 protein.
 GN RIORF151.
 OS Agrobacterium rhizogenes.
 OG Plasmid pRI1724.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Rhizobiaceae; Rhizobium.
 OX NCBI_TaxID=359;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MAFF03-01724;
 RA Moriguchi K., Maeda Y., Satou M., Satuti N., Kataoka M., Tanaka N.,
 RA Yoshida K.;
 RT "The complete nucleotide sequence of a Ri (root inducing) plasmid
 RT indicates its chimerical structure between Ti and Sym plasmids.";
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MAFF03-01724;
 RA Moriguchi K., Maeda Y., Satou M., Kataoka M., Tanaka N., Yoshida K.;
 RT "Analysis of unique variable region of a plant root inducing plasmid,
 RT pRI1724, by the construction of its physical map and library.";
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MAFF03-01724;
 RA Moriguchi K., Nishida T., Maeda Y., Tanaka N., Yoshida K.;
 RT "Genome structure of Ri plasmid (1): Construction of linking library
 RT and physical map of pRI1724 in Japanese Agrobacterium.";
 RL Nucleic Acids Symp. Ser. 39:189-190(1998).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MAFF03-01724;
 RX MEDLINE=20241294; PubMed=10780382;
 RA Maeda Y., Moriguchi K., Kataoka M., Satou M., Satuti N., Tanaka N.,
 RA Yoshida K.;
 RT "Genome structure of Ri plasmid (1): Sequencing analysis of T-DNA and
 RT its flanking regions of pRI1724 in Japanese Agrobacterium
 RT rhizogenes.";
 RL Nucleic Acids Symp. Ser. 42:67-68(1999).
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 DR EMBL: AP002086; BAB16270.1; -;
 DR InterPro: IPR001128; Cytochrome_P450.
 DR Pfam: PF00067; P450; 2.
 DR PRINTS: PR00385; P450.

```

DR PROSITE; PS00086; CYTOCHROME_P450; 1.
KW Heme; Monooxygenase; Oxidoreductase; Plasmid.
SQ SEQUENCE 450 AA; 51121 MW; B4C8FB5033084148 CRC64;

Query Match 43.6%; Score 48; DB 2; Length 450;
Best Local Similarity 66.7%; Pred. No. 31;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 7 EWTPKKPGRLRR 18
Db 73 EWRKRPFLRRQ 84
  ||| |||
  ||| |||

RESULT 13
Q9A9V4 PRELIMINARY; PRT; 735 AA.
ID Q9A9V4
AC Q9A9V4
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE GGDEF family protein.
GN CC0857.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacteriales;
OC Caulobacteraceae; Caulobacter.
OX NCBI_TaxID=155892;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19089 / CB15;
RX MEDLINE=21173698; PubMed=11259647;
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Uterback T., Tran K., Wolf A., Vanathavan J., Emoliaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus."
RT Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
DR EMBL; AE005762; AAK22842.1; -.
DR TIGR; CC0857; -.
DR InterPro; IPR001633; EAL.
DR InterPro; IPR000160; GGDEF.
DR Pfam; PF00563; EAL; 1.
DR Pfam; PF00990; GGDEF; 1.
DR SMART; SM00267; DUF1; 1.
DR SMART; SM00052; DUF2; 1.
DR TIGRFAMs; TIGR00254; GGDEF; 1.
KW Complete proteome.
SQ SEQUENCE 735 AA; 78705 MW; 9AF7E2E60F4201A2 CRC64;

Query Match 43.6%; Score 48; DB 16; Length 735;
Best Local Similarity 50.0%; Pred. No. 52;
Matches 10; Conservative 2; Mismatches 4; Indels 4; Gaps 1;

Qy 1 NNAGSV----EWTPKKPGRLR 16
Db 242 NHAGQVVATLDWTPPHQGR 261
  ||| |
  ||| |||

RESULT 14
Q9R6D1 PRELIMINARY; PRT; 391 AA.
ID Q9R6D1
AC Q9R6D1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE TIOF129 protein.
GN TIOF129.
OS Agrobacterium tumefaciens.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium.

```

```

OX NCBI_TaxID=358;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF301001;
RX MEDLINE=20184752; PubMed=10721727;
RA Suzuki K., Hattori Y., Uraji M., Ohta N., Iwata K., Murata K.,
RA Katoh A., Yoshida K.;
RT "Complete nucleotide sequence of a plant tumor-inducing Ti plasmid."
RL Gene 242:331-336(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF301001;
RX MEDLINE=98193120; PubMed=9524202;
RA Suzuki K., Ohta N., Hattori Y., Uraji M., Katoh A., Yoshida K.;
RT "Novel structural difference between nopaline- and octopline- type trbJ
RT gene: construction of genetic and physical map and sequencing of
RT trb/trai and rep gene clusters of a new Ti plasmid pTi-SAKURA."
RL Biochim. Biophys. Acta 1396:1-7(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF301001;
RA Hattori Y., Suzuki K., Ohta N., Uraji M., Katoh A., Yoshida K.;
RT "Genome structure of pTi-SAKURA(I): Strategy for DNA sequencing of a
RT Japanese cherry-Ti plasmid."
RL Nucleic Acids Symp. Ser. 37:159-160(1998).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF301001;
RA Ohta N., Suzuki K., Hattori Y., Uraji M., Katoh A., Yoshida K.;
RT "Genome structure of pTi-SAKURA (III): Characteristics of T-DNA."
RL Nucleic Acids Symp. Ser. 39:185-186(1998).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF301001;
RA Uraji M., Suzuki K., Ohta N., Hattori Y., Katoh A., Yoshida K.;
RT "Genome structure of pTi-SAKURA (IV): Characteristics of tra region."
RL Nucleic Acids Symp. Ser. 39:187-188(1998).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF301001;
RA Hattori Y., Suzuki K., Ohta N., Uraji M., Katoh A., Yoshida K.;
RT "Genome structure of pTi-SAKURA (V): Complete nucleotide sequence of
RT plasmid pTi-SAKURA's vir region in Agrobacterium tumefaciens."
RL Nucleic Acids Symp. Ser. 39:265-266(1998).
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; AB016260; BAA87754.1; -.
DR InterPro; IPR001128; Cytochrome_P450.
DR Pfam; PF00067; p450; 1.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
KW Heme; Monooxygenase; Oxidoreductase; Plasmid.
SQ SEQUENCE 391 AA; 44216 MW; 4012AD00FCDD396D CRC64;

Query Match 42.7%; Score 47; DB 2; Length 391;
Best Local Similarity 72.7%; Pred. No. 38;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 7 EWTPKKPGRLR 17
Db 42 EWRKRPFLRR 52
  ||| |||
  ||| |||

RESULT 15
O50242 PRELIMINARY; PRT; 419 AA.
ID O50242
AC O50242;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE P-450 monooxygenase.
GN VIRH1 OR ATU6150 OR AGR_PTI_272.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Plasmid pTiC58.

```

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Rhizobiaceae; Rhizobium.
 OX NCBI_TaxID=176299;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C58; PLASMID=PTIC58;
 RX MEDLINE=99009000; PubMed=9791116;
 RA Kalođeraki V.S., Winans S.C.;
 RT "Wound-released chemical signals may elicit multiple responses from an
 RT Agrobacterium tumefaciens strain containing an octopine-type Ti
 RT plasmid.";
 RL J. Bacteriol. 180:5660-5667(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC PLASMID=PTIC58;
 RX MEDLINE=21608550; PubMed=11743193;
 RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
 RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
 RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
 RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
 RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
 RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
 RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
 RA Gordon-Ramm B., Liang L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
 RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
 RA Nester E.W.;
 RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
 RT C58.";
 RL Science 294:2317-2323(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC PLASMID=PTIC58;
 RX MEDLINE=21608551; PubMed=11743194;
 RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
 RA Quorllo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullip L.,
 RA Houmel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
 RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
 RA Flanagan C., Crowell C., Gurson J., Lomo C., Sear C., Strub G.,
 RA Cielo C., Slater S.;
 RT "Genome sequence of the plant pathogen and biotechnology agent
 RT Agrobacterium tumefaciens C58.";
 RL Science 294:2323-2328(2001).
 CC -I- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 DR EMBL; AF034769; AAC1785.1; -;
 DR EMBL; AE009434; AAL46386.1; -;
 DR EMBL; AE007941; AAK91113.1; -;
 DR InterPro: IPR001128; Cytochrome_P450;
 DR Pfam: PF00067; P450; 1.
 DR PRINTS; PR00385; P450.
 DR PROSITE; PS00086; CYTOCHROME_P450; 1.
 KW Heme; Monooxygenase; Oxidoreductase; plasmid; Complete proteome.
 SQ SEQUENCE 419 AA; 47180 MW; EBAFB76BF99C13D CRC64;

Query Match 42.7%; Score 47; DB 16; Length 419;
 Best Local Similarity 72.7%; Pred. No. 41;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 EWTPKKPLRR 17
 |||||
 Db 42 EWRKRPFLRR 52

Search completed: August 27, 2003, 09:45:42
 Job time : 211 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 28, 2003, 13:47:55 ; Search time 125 Seconds
(without alignments)
146.242 Million cell updates/sec

Title: US-10-042-417A-91

Perfect score: 110
Sequence: 1 NNAGSVETPKPKPLRRROT 20

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 783778 seqs, 457005226 residues

Total number of hits satisfying chosen parameters: 1567556

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPTO.spool/US10042417/runat_19082003_141158_28005/app_query.fasta_1.199
-DB=Pending_Patents_NA_New -QFWT=fastp -SUFFIX=rpnp -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi
-LIST=45 -LOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10042417 -CGN_1_1_87 @runat_19082003_141158_28005 -NCPU=6 -ICPU=3
-NO_MMAP -LARGESQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Pending_Patents_NA_New.*
1: /cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq.*
2: /cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq.*
3: /cgn2_6/ptodata/1/pna/US07_NEW_COMB.seq.*
4: /cgn2_6/ptodata/1/pna/US08_NEW_COMB.seq.*
5: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq.*
6: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq.*
7: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	91	82.7	266	6	US-10-622-950-2658 Sequence 2658, Ap
2	91	82.7	2422	1	PCT-US03-14511A-10 Sequence 10, Appl
3	91	82.7	2422	7	US-60-490-890-716 Sequence 716, Appl
4	53.5	48.6	429	6	US-10-357-930-48311 Sequence 48311, A
5	52	47.3	2675	6	US-10-273-573-999 Sequence 999, App
6	52	47.3	48746	7	US-60-485-450-12024 Sequence 12024, A
7	50	45.5	2308	6	US-10-451-901-24 Sequence 24, Appl
8	50	45.5	2623	6	US-10-470-868-131 Sequence 131, App
9	50	45.5	5245	6	US-10-404-841-1 Sequence 1, Appl
c 10	50	45.5	17885	6	US-10-273-573-590 Sequence 590, App
c 11	49	44.5	457	6	US-10-357-930-14414 Sequence 14414, A

12	49	44.5	620	6	US-10-626-717-3447 Sequence 3447, Ap
13	49	44.5	1113	6	US-10-612-783-507 Sequence 507, App
14	49	44.5	152321	7	US-60-485-450-12128 Sequence 12128, A
c 15	49	44.5	2256646	6	US-10-470-565-1 Sequence 1, Appli
16	48	43.6	306	6	US-10-626-717-8670 Sequence 8670, Ap
c 17	48	43.6	1074	6	US-10-273-573-4718 Sequence 4718, Ap
18	48	43.6	1140	1	PCT-US03-10142-5 Sequence 5, Appli
19	48	43.6	2078	1	PCT-US03-10142-7 Sequence 7, Appli
c 20	48	43.6	2208	6	US-10-273-573-4705 Sequence 4705, Ap
c 21	48	43.6	2526	6	US-10-273-573-4263 Sequence 4263, Ap
c 22	48	43.6	3102	6	US-10-273-573-4266 Sequence 4266, Ap
c 23	48	43.6	8473	5	US-09-851-682B-2 Sequence 2, Appli
24	48	43.6	11736	6	US-10-291-265-218 Sequence 218, App
c 25	48	43.6	171096	6	US-10-612-869-8 Sequence 8, Appli
26	47.5	43.2	2582	6	US-10-374-780A-1791 Sequence 1791, Ap
c 27	47	42.7	201	7	US-60-487-610-44554 Sequence 44554, A
c 28	47	42.7	2032	6	US-10-273-573-3157 Sequence 3157, Ap
c 29	47	42.7	7509	6	US-10-600-058-4 Sequence 4, Appli
c 30	47	42.7	7602	7	US-60-478-196-159 Sequence 159, App
c 31	47	42.7	93405	7	US-60-487-610-19470 Sequence 19470, A
c 32	47	42.7	2256646	6	US-10-470-565-1 Sequence 1, Appli
c 33	46.5	42.3	2473	6	US-10-612-783-1948 Sequence 1948, Ap
c 34	46.5	42.3	15718	7	US-60-487-610-19608 Sequence 19608, A
c 35	46	41.8	454	6	US-10-626-717-7677 Sequence 7677, Ap
c 36	46	41.8	472	6	US-10-626-717-4297 Sequence 4297, Ap
c 37	46	41.8	635	6	US-10-298-226-74 Sequence 74, Appl
c 38	46	41.8	640	6	US-10-298-226-73 Sequence 73, Appl
c 39	46	41.8	1324	6	US-10-374-780A-1009 Sequence 1009, Ap
c 40	46	41.8	1656	1	PCT-US03-19156-1 Sequence 1, Appli
41	46	41.8	1656	6	US-10-464-811-1 Sequence 1, Appli
c 42	46	41.8	1866	6	US-10-293-244-2724 Sequence 2724, Ap
c 43	46	41.8	2880	1	PCT-US03-23245-70 Sequence 70, Appl
c 44	46	41.8	2916	6	US-10-293-244-756 Sequence 756, App
c 45	46	41.8	3074	6	US-10-341-434-23 Sequence 23, Appl

ALIGNMENTS

RESULT 1

US-10-622-950-2658
; Sequence 2658, Application US/10622950
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.; Delegeane, Angelo M.;
; APPLICANT: Stuart, Susan G.; Stuve, Laura L.;
; APPLICANT: Mullahy, Sara J.; Naughton, Rebecca E.
; TITLE OF INVENTION: MOLECULES FOR DISEASE DETECTION AND TREATMENT
; FILE REFERENCE: PD-1100 CIP
; CURRENT APPLICATION NUMBER: US/10/622,950
; PRIOR FILING DATE: 2003-02-17
; PRIOR APPLICATION NUMBER: 09/534,845
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 09/532,315
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 09/533,804
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 09/533,806
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 09/534,719
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 09/534,843
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 09/535,897
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 09/535,170
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 09/534,840
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 09/534,855
; Remaining Prior-Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 3122
; SOFTWARE: PERL Program
; SEQ ID NO 2658

```
; LENGTH: 266
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: hu00253345
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1) ... (266)
; OTHER INFORMATION: a, t, c, g, or other
US-10-622-950-2658

Alignment Scores:
Pred. No.: 3.43e-05 Length: 266
Score: 91.00 Matches: 18
Percent Similarity: 94.74% Conservative: 0
Best Local Similarity: 94.74% Mismatches: 1
Query Match: 82.73% Indels: 0
DB: 6 Gaps: 0

US-10-042-417A-91 (1-20) x US-10-622-950-2658 (1-266)
QY 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
Db 141 AATGCCGGTTCTGTGGAGCAGACGCCCAAGAAGCCTGGCCTCAGAAGACGTCAAACG 197

RESULT 2
PCT-US03-14511A-10
; Sequence 10, Application PC/TUS0314511A
; GENERAL INFORMATION:
; APPLICANT: Greenlee, Winner and Sullivan, P.C.
; TITLE OF INVENTION: Protein Complex Purification
; FILE REFERENCE: 53-02
; CURRENT APPLICATION NUMBER: PCT/US03/14511A
; CURRENT FILING DATE: 2003-05-09
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 2422
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US03-14511A-10

Alignment Scores:
Pred. No.: 0.000387 Length: 2422
Score: 91.00 Matches: 18
Percent Similarity: 94.74% Conservative: 0
Best Local Similarity: 94.74% Mismatches: 1
Query Match: 82.73% Indels: 0
DB: 1 Gaps: 0

US-10-042-417A-91 (1-20) x PCT-US03-14511A-10 (1-2422)
QY 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
Db 1003 AATGCCGGTTCTGTGGAGCAGACGCCCAAGAAGCCTGGCCTCAGAAGACGTCAAACG 1059

RESULT 3
US-60-490-890-716
; Sequence 716, Application US/60490890
; GENERAL INFORMATION:
; APPLICANT: Li, Martha
; APPLICANT: Rupnow, Brent A.
; APPLICANT: Webster, Kevin R.
; APPLICANT: Jackson, Donald
; APPLICANT: Wong, Tai W.
; TITLE OF INVENTION: BIOMARKERS OF CYCLIN-DEPENDENT KINASE MODULATION
; FILE REFERENCE: D0310 PSP
; CURRENT APPLICATION NUMBER: US/60/490,890
; CURRENT FILING DATE: 2003-07-29
; NUMBER OF SEQ ID NOS: 2779
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 716

; LENGTH: 2422
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-490-890-716

Alignment Scores:
Pred. No.: 0.000387 Length: 2422
Score: 91.00 Matches: 18
Percent Similarity: 94.74% Conservative: 0
Best Local Similarity: 94.74% Mismatches: 1
Query Match: 82.73% Indels: 0
DB: 1 Gaps: 0

US-10-042-417A-91 (1-20) x US-60-490-890-716 (1-2422)
QY 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
Db 1003 AATGCCGGTTCTGTGGAGCAGACGCCCAAGAAGCCTGGCCTCAGAAGACGTCAAACG 1059

RESULT 4
US-10-357-930-48311
; Sequence 48311, Application US/10357930
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 05/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 48311
; LENGTH: 429
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-357-930-48311

Alignment Scores:
Pred. No.: 20.2 Length: 429
Score: 53.50 Matches: 10
Percent Similarity: 61.54% Conservative: 6
Best Local Similarity: 38.46% Mismatches: 3
Query Match: 48.64% Indels: 7
DB: 6 Gaps: 1

US-10-042-417A-91 (1-20) x US-10-357-930-48311 (1-429)
QY 1 AsnAsnAlaGlySerValGluTrpThrProLys-----LysPro 13
Db 108 AACAGGAGGAGGCTGTGGAGTGGGCCCAAGAACGCTGCCAGGAGCGCTGAGCGCT 167

QY 14 GlyLeuArgArgGln 19
Db 168 GGGATGAACAGAGAGAG 185

RESULT 5
```

```

Query Match:      47.27%   Indels:      0
DB:              7       Gaps:         0

US-10-042-417A-91 (1-20) x US-60-485-450-12024 (1-48746)

Qy    3 AlaGlySerValGIutrrPhrProLysLysProGlyLeuArgArgGln 19
Db    43317 GCGGGTCTCCGGGAGTGGAAACCGGCCGACTCCGGGAACGAAGCACCGGC 43367

RESULT 7
US-10-451-901-24
; Sequence 24, Application US/10451901
; GENERAL INFORMATION:
; APPLICANT: INCYTE CORPORATION; BAUGHN, Mariah R.;
; APPLICANT: LU, Yan; ARVIZU, Chandra S.;
; APPLICANT: RAMKUMAR, Jayalaxmi; YAO, Monique G.;
; APPLICANT: POLICKY, Jennifer L.; WALIA, Nairender K.;
; APPLICANT: TRIBOULEY, Catherine M.; YUE, Henry;
; APPLICANT: BATRA, Sajeev; DING, Li;
; APPLICANT: LAI, Preeti G.; BOROWSKY, Mark L.;
; APPLICANT: LU, Dyung Aina M.; GANDHI, Ameena R.;
; APPLICANT: GRIFFIN, Jennifer A.; XU, Yuming;
; APPLICANT: AZIMZAI, Yalda; GIETZEN, Kimberly J.;
; APPLICANT: TANG, Y. Tom; WARREN, Bridget A.;
; APPLICANT: MASON, Patricia M.; BURFORD, Neil;
; APPLICANT: HAFALIA, April J.A.; LEE, Ernestine A.;
; APPLICANT: YANG, Junming; GORVAD, Ann E.;
; APPLICANT: EMERLING, Brooke M.; MARQUIS, Joseph P.;
; APPLICANT: LEE, Soo Yeun; SWARNAKAR, Anita;
; APPLICANT: REDDY, Roopa M.; JIANG, Xin;
; APPLICANT: JACKSON, Alan A.
; TITLE OF INVENTION: NUCLEIC ACID-ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0869 USN
; CURRENT APPLICATION NUMBER: US/10/451,901
; CURRENT FILING DATE: 2003-06-23
; PRIOR APPLICATION NUMBER: PCT/US01/50256
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: US 60/257,714
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: US 60/260,081
; PRIOR FILING DATE: 2001-01-05
; PRIOR APPLICATION NUMBER: US 60/262,302
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/263,823
; PRIOR FILING DATE: 2001-01-23
; PRIOR APPLICATION NUMBER: US 60/266,088
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: US 60/348,442
; PRIOR FILING DATE: 2001-10-29
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PERL Program
; SEQ ID NO 24
; LENGTH: 2308
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: incyte ID No: 6428773CB1
US-10-451-901-24

Alignment Scores:
Pred. No.:      421          Length:      2308
Score:          50.00        Matches:      7
Percent Similarity: 76.47%   Conservative: 6
Best Local Similarity: 41.18% Mismatches:     4
Query Match:    45.45%      Indels:       0
DB:             6           Gaps:         0

US-10-042-417A-91 (1-20) x US-10-451-901-24 (1-2308)

Qy    3 AlaGlySerValGIutrrPhrProLysLysProGlyLeuArgArgGln 19
Db    18 GCAGGAGCTCTTGACTGGCGCCGCCGCCGCTGGACTGAGCAAGCACC 68

```

```

Query Match: 45.45% Indels: 0
DB: Gaps: 0

US-10-042-417A-91 (1-20) x US-10-404-841-1 (1-5245)

Qy 1 AsnAsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArg 17
Db 2537 TCACTGCTCTAGCTCAGATGGAGTCTCTTCATACCCCGGCTCGAGAG 2487

RESULT 10
US-10-273-573-590
; Sequence 590, Application US/10273573
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-066
; CURRENT APPLICATION NUMBER: US/10/273, 573
; CURRENT FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: 09/522, 929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: 09/770, 160
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 10994
; SOFTWARE: Custom
; SEQ ID NO 590
; LENGTH: 17885
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (17859)..(15370)
; OTHER INFORMATION: 100% homologous to Homo sapiens acetyl LDL rece
; OTHER INFORMATION: number D86884, Smith-Waterman Score=4745.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(17885)
; OTHER INFORMATION: n = a,t,c or g
US-10-273-573-590

Alignment Scores:
Pred. No.: 3.98e+03 Length: 17885
Score: 50.00 Matches: 7
Percent Similarity: 78.57% Conservative: 4
Best Local Similarity: 50.00% Mismatches: 3
Query Match: 45.45% Indels: 0
DB: Gaps: 0

US-10-042-417A-91 (1-20) x US-10-273-573-590 (1-17885)

Qy 6 ValGluTrpThrProLysLysProGlyLeuArgArgGln 19
Db 17600 CTGCAGTGGGCCCAAGAAATCCAGGTTTCATCGACAGAG 17641

RESULT 11
US-10-357-930-14414/G
; Sequence 14414, Application US/10357930
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THI
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357, 930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785, 276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183, 319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189, 862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207, 454

```

;; PRIOR FILING DATE: 2000-05-25
;; PRIOR APPLICATION NUMBER: 60/211,314
;; PRIOR FILING DATE: 2000-06-09
;; PRIOR APPLICATION NUMBER: 60/219,007
;; PRIOR FILING DATE: 2000-07-18
;; PRIOR APPLICATION NUMBER: 60/255,281
;; PRIOR FILING DATE: 2000-12-13
;; NUMBER OF SEQ ID NOS: 62232
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 14414
;; LENGTH: 457
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-10-357-930-14414

Alignment Scores:
Pred. No.: 100 Length: 457
Score: 49.00 Matches: 11
Percent Similarity: 52.00% Conservative: 2
Best Local Similarity: 44.00% Mismatches: 2
Query Match: 44.55% Indels: 10
DB: 6 Gaps: 1

US-10-042-417A-91 (1-20) x US-10-357-930-14414 (1-457)

Qy 6 ValGluTrpThrProLys-----LysProGlyLeu 15
:::|||||
Db 324 ATGAATGGACCCCAAAATTAGGGTTTTCCTTGGGCCAAGGATTGTCCTGGATTA 265
Qy 16 ArgArgGlnThr 20
|||:::|
Db 264 AGGAAAGGGTGACA 250

RESULT 12

US-10-626-717-3447
;; Sequence 3447, Application US/10626717
;; GENERAL INFORMATION:
;; APPLICANT: Andersen, Scott E.
;; APPLICANT: Byrum, Joseph R.
;; APPLICANT: De La Pena, Robert C.
;; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
;; TITLE OF INVENTION: PLANTS
;; FILE REFERENCE: 38-21(15878)D
;; CURRENT APPLICATION NUMBER: US/10/626,717
;; CURRENT FILING DATE: 2003-07-25
;; PRIOR APPLICATION NUMBER: US 10/304,123
;; PRIOR FILING DATE: 2002-11-26
;; PRIOR APPLICATION NUMBER: US 09/594,596
;; PRIOR FILING DATE: 2000-06-15
;; NUMBER OF SEQ ID NOS: 10952
;; SEQ ID NO 3447
;; LENGTH: 620
;; TYPE: DNA
;; ORGANISM: Triticum aestivum
;; FEATURE:
;; NAME/KEY: unsure
;; LOCATION: (1)..(620)
;; OTHER INFORMATION: unsure at all n locations
;; FEATURE:
;; OTHER INFORMATION: Clone ID: UC-tsfl9316108c1lb1
US-10-626-717-3447

Alignment Scores:
Pred. No.: 140 Length: 620
Score: 49.00 Matches: 9
Percent Similarity: 61.11% Conservative: 2
Best Local Similarity: 50.00% Mismatches: 7
Query Match: 44.55% Indels: 0
DB: 6 Gaps: 0

US-10-042-417A-91 (1-20) x US-10-626-717-3447 (1-620)

Qy 3 AlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20

Db 272 GCAGGGCGGGGTTCTTGGACCCCTCGCTCCCTTAGAAGGCTGGAAGC 325
||||| ||||| ||| |||||
RESULT 13
US-10-612-783-507
;; Sequence 507, Application US/10612783
;; GENERAL INFORMATION:
;; APPLICANT: La Rosa, Thomas J.
;; APPLICANT: Kovalic, David K.
;; APPLICANT: Zhou, Yihua
;; APPLICANT: Cao, Yongwei
;; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
;; TITLE OF INVENTION: Plants
;; FILE REFERENCE: 38-21(53373)A
;; CURRENT APPLICATION NUMBER: US/10/612,783
;; CURRENT FILING DATE: 2003-07-02
;; NUMBER OF SEQ ID NOS: 7098
;; SEQ ID NO 507
;; LENGTH: 1113
;; TYPE: DNA
;; ORGANISM: Zea mays
;; FEATURE:
;; OTHER INFORMATION: Clone ID: PAT_MRT4577_101123C.1
US-10-612-783-507

Alignment Scores:
Pred. No.: 266 Length: 1113
Score: 49.00 Matches: 8
Percent Similarity: 71.43% Conservative: 2
Best Local Similarity: 57.14% Mismatches: 4
Query Match: 44.55% Indels: 0
DB: 6 Gaps: 0

US-10-042-417A-91 (1-20) x US-10-612-783-507 (1-1113)

Qy 6 ValGluTrpThrProLysLysProGlyLeuArgArgGln 19
||| ||| :::::|||||
Db 403 GTACGATGGCGGCACGCGCGGCGGACTACGCGTGCCCGAG 444

RESULT 14

US-60-485-450-12128
;; Sequence 12128, Application US/60485450
;; GENERAL INFORMATION:
;; APPLICANT: CARGILL, Michele
;; APPLICANT: CHANG, Sheng-Yung
;; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;; TITLE OF INVENTION: RESPONSE TO INTERPERON TREATMENT IN HEPATITIS C
;; TITLE OF INVENTION: VIRUS-INFECTED SUBJECTS, METHODS OF DETECTION AND USES
;; TITLE OF INVENTION: THEREOF
;; FILE REFERENCE: CL001470
;; CURRENT APPLICATION NUMBER: US/60/485,450
;; CURRENT FILING DATE: 2003-07-09
;; NUMBER OF SEQ ID NOS: 47859
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 12128
;; LENGTH: 152321
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: (1)..(152321)
;; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables
US-60-485-450-12128

Alignment Scores:
Pred. No.: 5,76e+04 Length: 152321
Score: 49.00 Matches: 10
Percent Similarity: 73.33% Conservative: 1
Best Local Similarity: 66.67% Mismatches: 4
Query Match: 44.55% Indels: 0
DB: 7 Gaps: 0

US-10-042-417A-91 (1-20) x US-60-485-450-12128 (1-152321)

QY 5 SerValGluTrpThrProLysLysProGlyLeuArgArgGln 19
 Db 96516 AGCTTGAATGGTAGCTTAAGAACCCAGGAGGAGGAGGAG 96560

RESULT 15

US-10-470-565-1/c
 ; Sequence 1, Application US/10470565
 ; GENERAL INFORMATION:
 ; APPLICANT: Societe des Produits Nestle S.A.
 ; TITLE OF INVENTION: NCC2705 - the genome of a Bifidobacterium
 ; FILE REFERENCE: 80290/WO
 ; CURRENT APPLICATION NUMBER: US/10/470,565
 ; CURRENT FILING DATE: 2003-07-29
 ; PRIOR APPLICATION NUMBER: EP 01102050.0
 ; PRIOR FILING DATE: 2001-01-30
 ; NUMBER OF SEQ ID NOS: 2
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 1
 ; LENGTH: 2256646
 ; TYPE: DNA
 ; ORGANISM: Bifidobacterium longum
 US-10-470-565-1

Alignment Scores:
 Pred. No.: 8.05e+05 Length: 2256646
 Score: 49.00 Matches: 8
 Percent Similarity: 70.59% Conservative: 4
 Best Local Similarity: 47.06% Mismatches: 5
 Query Match: 44.55% Indels: 0
 DB: 6 Gaps: 0

US-10-042-417a-91 (1-20) x US-10-470-565-1 (1-2256646)

QY 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArg 18
 Db 352724 AGCGCATCTTCGCTGTGGACACCCGAGCCGAGCATGGCAGCGGA 352674

Search completed: August 28, 2003, 16:27:17
 Job time : 432 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2003, 09:28:35 ; Search time 77 Seconds
(without alignments)
24.979 Million cell updates/sec

Title: US-10-042-417A-91

Perfect score: 110

Sequence: 1 NAGSVEWTPKKPGLRRROT 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_76.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	91	82.7	198	2 I52718	gene p27Kip1 prote
2	74	67.3	197	2 I49064	cyclin-cdk inhibit
3	54	49.1	874	2 T15570	hypothetical prote
4	48	43.6	735	2 F87355	GDEF family prote
5	47	42.7	419	2 A03246	P-450 monooxygenas
6	47	42.7	422	1 A32306	cytochrome P450 cy
7	46	41.8	303	2 A12543	hypothetical prote
8	45.5	41.4	710	2 S26006	hypothetical prote
9	45	40.9	602	2 T47794	hypothetical prote
10	45	40.9	951	2 T00017	gene ADAMTS-1 prot
11	44.5	40.5	493	2 D27255	hypothetical prote
12	44	40.0	126	2 G69132	ribosomal protein
13	44	40.0	129	2 G69313	SSU ribosomal prot
14	44	40.0	167	2 B75545	conserved hypothet
15	44	40.0	262	2 T37381	hypothetical prote
16	44	40.0	458	2 T33814	hypothetical prote
17	44	40.0	683	2 H87912	protein B0205.1 [i
18	44	40.0	796	2 G25582	probable acetyl-Co
19	43.5	39.5	2054	2 T32413	bifunctional penic
20	43	39.1	284	2 H89962	conserved hypothet
21	43	39.1	603	1 Q0B0D1	HHLF5 protein - hu
22	43	39.1	801	2 T14338	sucrose synthase (
23	43	39.1	805	1 YUPOS	sucrose synthase (
24	43	39.1	807	2 T19836	hypothetical prote
25	43	39.1	808	2 S37560	sucrose synthase (
26	42.5	38.6	311	2 H70829	hypothetical prote
27	42	38.2	235	2 F89472	protein ZC53.7 [im
28	42	38.2	302	1 WMBE1E	secretory compleme
29	42	38.2	307	2 B82394	hypothetical prote

30	42	38.2	309	2 F83044	nitrate-inducible
31	42	38.2	333	2 D70792	hypothetical prote
32	42	38.2	343	2 T02399	hypothetical prote
33	42	38.2	360	1 WMBE2E	membrane-bound com
34	42	38.2	458	2 T34574	hypothetical prote
35	42	38.2	470	2 G70564	probable membrane
36	42	38.2	472	2 T29313	hypothetical prote
37	42	38.2	532	2 A65138	hypothetical 60.3
38	42	38.2	532	2 A91162	probable 2-compone
39	42	38.2	532	2 B86008	probable 2-compone
40	42	38.2	657	2 T42191	nixB protein - Esc
41	42	38.2	695	2 S66662	protein-glutamine
42	42	38.2	943	2 JC4081	sucrase/fructanase
43	42	38.2	958	2 C86308	F20D23.9 protein -
44	42	38.2	1297	2 T52065	probable myb-relat
45	42	38.2	1331	2 AE1843	hypothetical prote

ALIGNMENTS

RESULT 1

I52718
gene p27Kip1 protein - human
C:Species: Homo.sapiens (man)
C:Date: 01-Nov-1996 #sequence_revision 01-Nov-1996 #text_change 21-Jul-2000
C:Accession: I52718
R:Pietenpol, J.A.; Bohlander, S.K.; Sato, Y.; Papadopoulos, N.; Liu, B.; Friedman, C.
Cancer Res. 55, 1206-1210, 1995
A:Title: Assignment of the human p27Kip1 gene to 12p13 and its analysis in leukemias.
A:Reference number: I52718; MUID:95188144; PMID:7882309
A:Accession: I52718
A>Status: preliminary, translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-198 <RES>
A:Cross-references: GI:S76988; NID:998402; PIDN:AAD14244.1; PID:g4261944
C:Genetics:
A:Gene: p27Kip1
A:Introns: 159/1

Query Match 82.7%; Score 91; DB 2; Length 198;
Best Local Similarity 94.7%; Pred. No. 4e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 NAGSVEWTPKKPGLRRROT 20
||||| |||||||||
DB 180 NAGSVEQTPKKPGLRRROT 198

RESULT 2

I49064
cyclin-cdk inhibitor p27 - mouse
C:Alternate names: CDI p27; G1 cyclin-cyclin-dependent kinase inhibitor p27
C:Species: Mus musculus (house mouse)
C:Date: 09-Mar-1996 #sequence_revision 09-Mar-1996 #text_change 05-Nov-1999
C:Accession: I49064
R:Toyoshima, H.; Hunter, T.
Cell 78, 67-74, 1994
A:Title: p27, a novel inhibitor of G1 cyclin-Cdk protein kinase activity, is related
A:Reference number: A54839; MUID:94306519; PMID:8033213
A:Accession: I49064
A:Molecule type: mRNA
A>Status: preliminary
A:Residues: 1-197 <RES>
A:Cross-references: EMBL:U10440; NID:9532771; PIDN:AAA21149.1; PID:9532772
C:Keywords: cell cycle control

Query Match 67.3%; Score 74; DB 2; Length 197;
Best Local Similarity 82.4%; Pred. No. 0.0013;
Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 NAGSVEWTPKKPGLRRR 18
||||| |||||||||

A:Residues: 1-303 <KUR>
 A:Cross-references: GB:AP003602; PIDN:BA877279.1; PID:g17134721; GSPDB:GN00181
 A:Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: alr7636
 A:Genome: plasmid

Query Match 41.8%; Score 46; DB 2; Length 303;
 Best Local Similarity 64.3%; Pred. No. 27;
 Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 NNAGSVETPKKPG 14
 | ||| ||| ||
 Db 269 NQTSVEFTPKYP 282

RESULT 8
 S26006
 hypothetical protein - liverwort (Marchantia polymorpha) mitochondrion
 C:Species: mitochondrion Marchantia polymorpha
 C:Date: 07-May-1993 #sequence_revision 07-May-1993 #text_change 21-Jul-2000
 C:Accession: S26006
 R:Oda, K.; Yamato, K.; Ohta, E.; Nakamura, Y.; Takemura, M.; Nozato, N.; Akashi, K.; Kan
 J. Mol. Biol. 223, 1-7, 1992
 A:Title: Gene organization deduced from the complete sequence of liverwort Marchantia po
 A:Reference number: S25941; MUID:92114051; PMID:1731062
 A:Accession: S26006
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-710 <ODA>
 A:Cross-references: EMBL:M68929; NID:g786182; PIDN:AAC09460.1; PID:g786246
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1992
 C:Genetics:
 A:Genome: mitochondrion
 C:Keywords: mitochondrion

Query Match 41.4%; Score 45.5; DB 2; Length 710;
 Best Local Similarity 45.5%; Pred. No. 72;
 Matches 10; Conservative 2; Mismatches 3; Indels 7; Gaps 1;

Qy 2 NAGSVETPKK-----PGLR 16
 | ||| ||| || || ||
 Db 129 NEGSPQWTPKRLVPRCOTPGDR 150

RESULT 9
 T47794
 hypothetical protein F17J16.170 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
 C:Accession: T47794
 R:D'Angelo, M.; Vezzi, A.; Modesto, D.; Pigazzi, M.; Valle, G.; Mewes, H.W.; Rudd, S.; I
 submitted to the Protein Sequence Database, April 2000
 A:Reference number: Z24476
 A:Accession: T47794
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-602 <DAN>
 A:Cross-references: EMBL:AL163527
 A:Experimental source: cultivar Columbia; BAC clone F17J16
 C:Genetics:
 A:Map position: 3
 A:Introns: 580/3
 A:Note: F17J16.170

Query Match 40.9%; Score 45; DB 2; Length 602;
 Best Local Similarity 50.0%; pred. No. 73;
 Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 2 NAGSVETPKKPG 13
 | | || ||| |||
 Db 279 NGGELEWAEPP 290

RESULT 10
 T00017
 gene ADAMTS-1 protein - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 21-Jul-2000
 C:Accession: T00017
 R:Kuno, K.; Lizasa, H.; Ohno, S.; Matsushima, K.
 Genomics 46, 466-471, 1997
 A:Title: The exon/intron organization and chromosomal mapping of the mouse ADAMTS-1 g
 A:Reference number: Z14055; MUID:98110583; PMID:9441751
 A:Accession: T00017
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-951 <KUN>
 A:Cross-references: EMBL:AB001735; NID:g2809056; PIDN:BAA24501.1; PID:g2809057
 A:Experimental source: strain 129SVJ
 C:Genetics:

A:Gene: ADAMTS-1
 A:Introns: 228/1; 343/3; 388/1; 444/1; 539/3; 602/1; 660/3; 719/2
 C:Superfamily: thrombospondin type 1 repeat homology
 F:542-598/Domain: thrombospondin type 1 repeat homology <THR3>

Query Match 40.9%; Score 45; DB 2; Length 951;
 Best Local Similarity 57.1%; Pred. No. 1.1e+02;
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 2 NAGSVETPKKPG 15
 | : || || || ||
 Db 623 NEPTVEWTPKYAGV 636

RESULT 11
 D72755
 hypothetical protein APE0033 - Aeropyrum pernix (strain K1)
 C:Species: Aeropyrum pernix
 C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
 C:Accession: D72755
 R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Ta
 awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.
 DNA Res. 6, 83-101, 1999
 A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aero
 A:Reference number: A72450; MUID:99310339; PMID:10382966
 A:Accession: D72755
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-493 <KAW>
 A:Cross-references: DBJ:AP000058; NID:g5103388; PIDN:BAA78942.1; PID:g5103421
 A:Experimental source: strain K1
 C:Genetics:
 A:Gene: APE0033
 C:Superfamily: phosphoenolpyruvate carboxykinase (ATP)

Query Match 40.5%; Score 44.5; DB 2; Length 493;
 Best Local Similarity 42.1%; Pred. No. 72;
 Matches 8; Conservative 5; Mismatches 5; Indels 1; Gaps 1;

Qy 2 NAGSVETPKKPG 20
 :: || | || || || ||
 Db 18 SSADIEWNP-PPGLLRRES 35

RESULT 12
 G69132
 ribosomal protein S6 - Methanobacterium thermoautotrophicum (strain Delta H)
 C:Species: Methanobacterium thermoautotrophicum
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 13-Aug-1999
 C:Accession: G69132
 R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T
 ; Qiu, D.; Spadarora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiواني,
 ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
 J. Bacteriol. 179, 7135-7155, 1997
 A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: fu

A:Reference number: A69000; MUID:98037514; PMID:9371463
A:Accession: G69132
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-126 <MTH>
A:Cross-references: GB:AE000812; GB:AE000666; NID:g2621298; PIDN:AAB84766.1; PID:g262131
A:Experimental source: strain Delta H
C:Genetics:
A:Gene: MTH260
A:Start codon: TTG
C:Superfamily: Haloarcula ribosomal protein H513

Query Match 40.0%; Score 44; DB 2; Length 126;
Best Local Similarity 44.4%; Pred. No. 23;
Matches 8; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 3 AGSVETPKPKGLRRRT 20
|||: |||: |||: |||:
Db 75 SGGVGYKPRRDRRRRT 92

RESULT 13
G69313
SSU ribosomal protein S6E (rps6E) homolog - Archaeoglobus fulgidus
C:Species: Archaeoglobus fulgidus
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 13-Aug-1999
C:Accession: G69313
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artlich, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Woese, C.R.; Venter, J.C.
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae
A:Reference number: A69250; MUID:98049343; PMID:9389475
A:Accession: G69313
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-129 <KLE>
A:Cross-references: GB:AE001068; GB:AE000782; NID:g2689391; PIDN:AAB90719.1; PID:g265010
C:Superfamily: Haloarcula ribosomal protein H513

Query Match 40.0%; Score 44; DB 2; Length 129;
Best Local Similarity 47.1%; Pred. No. 24;
Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 3 AGSVETPKPKGLRRRT 19
|||: |||: |||: |||:
Db 78 SGGVGYKPKGLRRRT 94

RESULT 14
B75545
conserved hypothetical protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 19-May-2000
C:Accession: B75545
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: B75545
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-167 <WHI>
A:Cross-references: GB:AE001884; GB:AE000513; NID:g6457890; PIDN:AAF09807.1; PID:g645789
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR0221
A:Map position: 1
C:Superfamily: hypothetical protein HI0925

Query Match 40.0%; Score 44; DB 2; Length 167;
Best Local Similarity 53.3%; Pred. No. 30;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 3 AGSVETPKPKGLRR 17
|||: |||: |||: |||:
Db 21 AGGLVWTPPHPLLRK 35

RESULT 15
T27381
hypothetical protein Y75B8A.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Feb-2000
C:Accession: T27381
R:Barlow, K.
submitted to the EMBL Data Library, November 1998
A:Reference number: Z20361
A:Accession: T27381
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-262 <WIL>
A:Cross-references: EMBL:AL033514; NID:el343251; PIDN:CAA22081.1; CESP:Y75B8A.1
A:Experimental source: clone Y75B8A
C:Genetics:
A:Gene: CESP:Y75B8A.1
A:Introns: 102/2; 173/1; 235/3
C:Superfamily: unassigned homeobox proteins; homeobox homology
C:Keywords: DNA binding; homeobox; nucleus; transcription regulation

Query Match 40.0%; Score 44; DB 2; Length 262;
Best Local Similarity 31.6%; Pred. No. 47;
Matches 6; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

QY 1 NNAGSVETPKPKGLRRRQ 19
|||: |||: |||: |||:
Db 172 NDGGQLEWTSSSHAMRKR 190

Search completed: August 27, 2003, 09:47:12
Job time: 82 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 28, 2003, 13:42:56 ; Search time 3669 Seconds
(without alignments)
178.620 Million cell updates/sec

Title: US-10-042-417A-91

Perfect score: 110

Sequence: 1 NNAGSVEWTPKPKPLRRROT 20

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 32997241 seqs, 16383922548 residues

Total number of hits satisfying chosen parameters: 65994482

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10042417/runat_19082003_141157_27972/app_query.fasta.1.199
-DB=Pending_Patents_NA_Main -QFWT=fastap -SUFFIX=rnnp -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=bloms62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10042417 -CGN_1.1_3868 -runat_19082003_141157_27972 -NCPU=6 -ICPU=3
-NO_MMAP -LARGESQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Pending_Patents_NA_Main.*

1: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq.*
2: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq.old.*
3: /cgn2_6/ptodata/2/pna/US07_COMB.seq.*
4: /cgn2_6/ptodata/2/pna/US07_COMB.seq.*
5: /cgn2_6/ptodata/2/pna/US08_COMB.seq.*
6: /cgn2_6/ptodata/2/pna/US081_COMB.seq.*
7: /cgn2_6/ptodata/2/pna/US082_COMB.seq.*
8: /cgn2_6/ptodata/2/pna/US083_COMB.seq.*
9: /cgn2_6/ptodata/2/pna/US084_COMB.seq.*
10: /cgn2_6/ptodata/2/pna/US085_COMB.seq.*
11: /cgn2_6/ptodata/2/pna/US086_COMB.seq.*
12: /cgn2_6/ptodata/2/pna/US087_COMB.seq.*
13: /cgn2_6/ptodata/2/pna/US088_COMB.seq.*
14: /cgn2_6/ptodata/2/pna/US089_COMB.seq.*
15: /cgn2_6/ptodata/2/pna/US090_COMB.seq.*
16: /cgn2_6/ptodata/2/pna/US091_COMB.seq.*
17: /cgn2_6/ptodata/2/pna/US092A_COMB.seq.*
18: /cgn2_6/ptodata/2/pna/US092B_COMB.seq.*
19: /cgn2_6/ptodata/2/pna/US093A_COMB.seq.*
20: /cgn2_6/ptodata/2/pna/US093B_COMB.seq.*
21: /cgn2_6/ptodata/2/pna/US094_COMB.seq.*
22: /cgn2_6/ptodata/2/pna/US095A_COMB.seq.*
23: /cgn2_6/ptodata/2/pna/US095B_COMB.seq.*
24: /cgn2_6/ptodata/2/pna/US095C_COMB.seq.*
25: /cgn2_6/ptodata/2/pna/US095D_COMB.seq.*
26: /cgn2_6/ptodata/2/pna/US096A_COMB.seq.*
27: /cgn2_6/ptodata/2/pna/US096B_COMB.seq.*
28: /cgn2_6/ptodata/2/pna/US096C_COMB.seq.*
29: /cgn2_6/ptodata/2/pna/US096D_COMB.seq.*
30: /cgn2_6/ptodata/2/pna/US096E_COMB.seq.*
31: /cgn2_6/ptodata/2/pna/US097A_COMB.seq.*
32: /cgn2_6/ptodata/2/pna/US097B_COMB.seq.*
33: /cgn2_6/ptodata/2/pna/US097C_COMB.seq.*
34: /cgn2_6/ptodata/2/pna/US098A_COMB.seq.*
35: /cgn2_6/ptodata/2/pna/US098B_COMB.seq.*
36: /cgn2_6/ptodata/2/pna/US098C_COMB.seq.*
37: /cgn2_6/ptodata/2/pna/US098D_COMB.seq.*
38: /cgn2_6/ptodata/2/pna/US099A_COMB.seq.*
39: /cgn2_6/ptodata/2/pna/US099B_COMB.seq.*
40: /cgn2_6/ptodata/2/pna/US099C_COMB.seq.*
41: /cgn2_6/ptodata/2/pna/US099D_COMB.seq.*
42: /cgn2_6/ptodata/2/pna/US099E_COMB.seq.*
43: /cgn2_6/ptodata/2/pna/US099F_COMB.seq.*
44: /cgn2_6/ptodata/2/pna/US100A_COMB.seq.*
45: /cgn2_6/ptodata/2/pna/US100B_COMB.seq.*
46: /cgn2_6/ptodata/2/pna/US101A_COMB.seq.*
47: /cgn2_6/ptodata/2/pna/US101B_COMB.seq.*
48: /cgn2_6/ptodata/2/pna/US102A_COMB.seq.*
49: /cgn2_6/ptodata/2/pna/US102B_COMB.seq.*
50: /cgn2_6/ptodata/2/pna/US103A_COMB.seq.*
51: /cgn2_6/ptodata/2/pna/US103B_COMB.seq.*
52: /cgn2_6/ptodata/2/pna/US104A_COMB.seq.*
53: /cgn2_6/ptodata/2/pna/US104B_COMB.seq.*
54: /cgn2_6/ptodata/2/pna/US6000_COMB.seq.*
55: /cgn2_6/ptodata/2/pna/US6001_COMB.seq.*
56: /cgn2_6/ptodata/2/pna/US6002_COMB.seq.*
57: /cgn2_6/ptodata/2/pna/US6003_COMB.seq.*
58: /cgn2_6/ptodata/2/pna/US6005_COMB.seq.*
59: /cgn2_6/ptodata/2/pna/US6006_COMB.seq.*
60: /cgn2_6/ptodata/2/pna/US6007_COMB.seq.*
61: /cgn2_6/ptodata/2/pna/US6008_COMB.seq.*
62: /cgn2_6/ptodata/2/pna/US6009_COMB.seq.*
63: /cgn2_6/ptodata/2/pna/US6010_COMB.seq.*
64: /cgn2_6/ptodata/2/pna/US6011_COMB.seq.*
65: /cgn2_6/ptodata/2/pna/US6012_COMB.seq.*
66: /cgn2_6/ptodata/2/pna/US6013_COMB.seq.*
67: /cgn2_6/ptodata/2/pna/US6014_COMB.seq.*
68: /cgn2_6/ptodata/2/pna/US6015_COMB.seq.*
69: /cgn2_6/ptodata/2/pna/US6016_COMB.seq.*
70: /cgn2_6/ptodata/2/pna/US6017_COMB.seq.*
71: /cgn2_6/ptodata/2/pna/US6018_COMB.seq.*
72: /cgn2_6/ptodata/2/pna/US6019_COMB.seq.*
73: /cgn2_6/ptodata/2/pna/US6020_COMB.seq.*
74: /cgn2_6/ptodata/2/pna/US6021_COMB.seq.*
75: /cgn2_6/ptodata/2/pna/US6022_COMB.seq.*
76: /cgn2_6/ptodata/2/pna/US6023A_COMB.seq.*
77: /cgn2_6/ptodata/2/pna/US6023B_COMB.seq.*
78: /cgn2_6/ptodata/2/pna/US6024_COMB.seq.*
79: /cgn2_6/ptodata/2/pna/US6025_COMB.seq.*
80: /cgn2_6/ptodata/2/pna/US6030_COMB.seq.*
81: /cgn2_6/ptodata/2/pna/US6031_COMB.seq.*
82: /cgn2_6/ptodata/2/pna/US6032_COMB.seq.*
83: /cgn2_6/ptodata/2/pna/US6033_COMB.seq.*
84: /cgn2_6/ptodata/2/pna/US6029_COMB.seq.*
85: /cgn2_6/ptodata/2/pna/US6030_COMB.seq.*
86: /cgn2_6/ptodata/2/pna/US6031_COMB.seq.*
87: /cgn2_6/ptodata/2/pna/US6032_COMB.seq.*
88: /cgn2_6/ptodata/2/pna/US6033_COMB.seq.*
89: /cgn2_6/ptodata/2/pna/US6034_COMB.seq.*
90: /cgn2_6/ptodata/2/pna/US6035_COMB.seq.*
91: /cgn2_6/ptodata/2/pna/US6036_COMB.seq.*
92: /cgn2_6/ptodata/2/pna/US6037_COMB.seq.*
93: /cgn2_6/ptodata/2/pna/US6038_COMB.seq.*
94: /cgn2_6/ptodata/2/pna/US6039_COMB.seq.*
95: /cgn2_6/ptodata/2/pna/US6040_COMB.seq.*
96: /cgn2_6/ptodata/2/pna/US6041_COMB.seq.*
97: /cgn2_6/ptodata/2/pna/US6042_COMB.seq.*
98: /cgn2_6/ptodata/2/pna/US6043_COMB.seq.*
99: /cgn2_6/ptodata/2/pna/US6044_COMB.seq.*
100: /cgn2_6/ptodata/2/pna/US6045_COMB.seq.*
101: /cgn2_6/ptodata/2/pna/US6046_COMB.seq.*

102: /cgn2_6/ptodata/2/pna/US6047_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	91	82.7	266	14	Sequence 1534, Ap
2	91	82.7	266	52	Sequence 2658, Ap
3	91	82.7	266	56	Sequence 1534, Ap
4	91	82.7	275	6	Sequence 9866, Ap
5	91	82.7	275	6	Sequence 9866, Ap
6	91	82.7	275	6	Sequence 9866, Ap
7	91	82.7	275	35	Sequence 9866, Ap
8	91	82.7	275	38	Sequence 9866, Ap
9	91	82.7	364	80	Sequence 115347, Ap
10	91	82.7	508	20	Sequence 5665, Ap
11	91	82.7	508	34	Sequence 3977, Ap
12	91	82.7	562	32	Sequence 3977, Ap
13	91	82.7	562	32	Sequence 18917, A
14	91	82.7	576	44	Sequence 18917, A
15	91	82.7	580	23	Sequence 12066, A
16	91	82.7	596	12	Sequence 34024, A
17	91	82.7	596	12	Sequence 1, Appli
18	91	82.7	596	36	Sequence 1, Appli
19	91	82.7	597	13	Sequence 1, Appli
20	91	82.7	597	13	Sequence 1, Appli
21	91	82.7	597	13	Sequence 1, Appli
22	91	82.7	597	15	Sequence 1, Appli
23	91	82.7	597	15	Sequence 1431, Ap
24	91	82.7	597	20	Sequence 5, Appli
25	91	82.7	597	21	Sequence 5, Appli
26	91	82.7	597	26	Sequence 8, Appli
27	91	82.7	597	31	Sequence 5, Appli
28	91	82.7	597	36	Sequence 5, Appli
29	91	82.7	597	36	Sequence 1, Appli
30	91	82.7	597	42	Sequence 1, Appli
31	91	82.7	597	47	Sequence 304, App
32	91	82.7	597	53	Sequence 8, Appli
33	91	82.7	670	32	Sequence 283, App
34	91	82.7	670	48	Sequence 283, App
35	91	82.7	1015	72	Sequence 976, App
36	91	82.7	1098	31	Sequence 7, Appli
37	91	82.7	1143	31	Sequence 5, Appli
38	91	82.7	1262	28	Sequence 10129, A
39	91	82.7	1262	28	Sequence 3034, Ap
40	91	82.7	1266	87	Sequence 12507, A
41	91	82.7	1420	31	Sequence 1, Appli
42	91	82.7	1615	47	Sequence 37574, A
43	91	82.7	1615	99	Sequence 588, App
44	91	82.7	1615	100	Sequence 9230, Ap
45	91	82.7	1615	100	Sequence 2744, Ap

ALIGNMENTS

RESULT 1
US-08-943-980-1534
; Sequence 1534, Application US/08943980
; GENERAL INFORMATION:
; APPLICANT: Gooding, Douglas H.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Akerblom, Ingrid E.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Naughton, Rebecca E.
; APPLICANT: Klingler, Tod M.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM

; TITLE OF INVENTION: HUMAN LEUKOCYTES
; NUMBER OF SEQUENCES: 3973
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA: US/08/943,980
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/027,838
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: CERRONE, MICHAEL C., Ph.D.
; REGISTRATION NUMBER: 39132
; REFERENCE/DOCKET NUMBER: PD-0246P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 855-0555
; TELEFAX: (415) 845-4166
; INFORMATION FOR SEQ ID NO: 1534:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 266 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; IMMEDIATE SOURCE:
; CLONE: 1876174
; US-08-943-980-1534

Alignment Scores:
Pred. No.: 0.00236 Length: 266
Score: 91.00 Matches: 18
Percent Similarity: 94.74% Conservative: 0
Best Local Similarity: 94.74% Mismatches: 1
Query Match: 82.73% Indels: 0
DB: 14 Gaps: 0

US-10-042-417a-91 (1-20) x US-08-943-980-1534 (1-266)
QY 2 AsnAlaGlySerValGluTrpThrProLysProGlyLeuArgArgGlnThr 20
Db 141 AATGCCGGTCTCTGTGGAGCAGACGCCCAAGACCTGGCTCAGAACGCTCAACG 197

RESULT 2
US-10-406-815-2658
; Sequence 2658, Application US/10406815
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.; Delegeane, Angelo M.;
; APPLICANT: Stuart, Susan G.; Stuve, Laura L.;
; APPLICANT: Mullahy, Sara J.; Naughton, Rebecca E.
; TITLE OF INVENTION: MOLECULES FOR DISEASE DETECTION AND TREATMENT
; FILE REFERENCE: PD-1100 CIP
; CURRENT APPLICATION NUMBER: US/10/406,815
; CURRENT FILING DATE: 2003-04-02
; PRIOR APPLICATION NUMBER: 09/534,845
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 09/532,315
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 09/533,804
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 09/533,806
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 09/534,719

;; PRIOR FILING DATE: 2000-03-24
;; PRIOR APPLICATION NUMBER: 09/534,843
;; PRIOR FILING DATE: 2000-03-24
;; PRIOR APPLICATION NUMBER: 09/535,897
;; PRIOR FILING DATE: 2000-03-24
;; PRIOR APPLICATION NUMBER: 09/535,170
;; PRIOR FILING DATE: 2000-03-24
;; PRIOR APPLICATION NUMBER: 09/534,840
;; PRIOR FILING DATE: 2000-03-24
;; PRIOR APPLICATION NUMBER: 09/534,855
;; PRIOR FILING DATE: 2000-03-24
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 3122
;; SOFTWARE: PERL Program
;; SEQ ID NO 2658
;; LENGTH: 266
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc feature
;; OTHER INFORMATION: Incyte ID No: hu00253345
;; FEATURE:
;; NAME/KEY: unsure
;; LOCATION: (1) ... (266)
;; OTHER INFORMATION: a, t, c, g, or other
US-10-406-815-2658

Alignment Scores:
Pred. No.: 0.00236 Length: 266
Score: 91.00 Matches: 18
Percent Similarity: 94.74% Conservative: 0
Best Local Similarity: 94.74% Mismatches: 1
Query Match: 82.73% Indels: 0
DB: 52 Gaps: 0

US-10-042-417A-91 (1-20) x US-10-406-815-2658 (1-266)

Qy 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
Db 141 AATCGCGTTCTGTGGAGCAGACGCCCAAGAGCGCTGGCCTCAGAAAGCGTCAACG 197

RESULT 3
US-027-938-1534
; Sequence 1534, Application US/60027838
; GENERAL INFORMATION:
; APPLICANT: Gooding, Douglas H.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Akerblom, Ingrid E.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Naughton, Rebecca E.
; APPLICANT: Klingler, Tod M.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
; TITLE OF INVENTION: HUMAN LEUKOCYTES
; NUMBER OF SEQUENCES: 3973
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/60/027,838
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:

;; NAME: CERRONE, MICHAEL C., Ph.D.
;; REGISTRATION NUMBER: 39132
;; REFERENCE/DOCKIT NUMBER: PD-0246P
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 855-0555
;; TELEFAX: (415) 845-4166
;; INFORMATION FOR SEQ ID NO: 1534:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 266 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: cDNA
;; IMMEDIATE SOURCE:
;; CLONE: 1876174
US-60-027-838-1534

Alignment Scores:
Pred. No.: 0.00236 Length: 266
Score: 91.00 Matches: 18
Percent Similarity: 94.74% Conservative: 0
Best Local Similarity: 94.74% Mismatches: 1
Query Match: 82.73% Indels: 0
DB: 56 Gaps: 0

US-10-042-417A-91 (1-20) x US-60-027-838-1534 (1-266)

Qy 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
Db 141 AATCGCGTTCTGTGGAGCAGACGCCCAAGAGCGCTGGCCTCAGAAAGCGTCAACG 197

RESULT 4
US-08-196-363-9866
; Sequence 9866, Application US/08196363
; GENERAL INFORMATION:
; APPLICANT: Kerlevage, Anthony
; APPLICANT: Fraser, Claire M.
; APPLICANT: Kirkness, Ewen
; APPLICANT: Sutclon, Granger
; APPLICANT: Gocayne, Jeannine
; APPLICANT: Liu, Li-ing
; APPLICANT: Fitzgerald, Lisa
; APPLICANT: Adams, Mark
; APPLICANT: Lee, Normal
; APPLICANT: Fuldner, Rebecca
; APPLICANT: Fleischmann, Robert
; APPLICANT: Bult, Carol
; APPLICANT: Blake, Judy
; APPLICANT: White, Owen
; APPLICANT: Clayton, Rebecca
; APPLICANT: Pellegrino, Susan
; TITLE OF INVENTION: Human Genes, Sequences, and Expression
; TITLE OF INVENTION: Products
; NUMBER OF SEQUENCES: 16818
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi, Stewart, &
; ADDRESSEE: Olstein
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 5.0
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/196,363
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:

```
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Olstein, Elliot
; REGISTRATION NUMBER: 24,025
; REFERENCE/DOCKET NUMBER: 325800-91
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (201) 994-1700
; TELEFAX: (201) 994-1744
; INFORMATION FOR SEQ ID NO: 9866:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 275 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-196-363-9866

Alignment Scores:
Pred. No.: 0.00244 Length: 275
Score: 91.00 Matches: 18
Percent Similarity: 94.74% Conservatve: 0
Best Local Similarity: 94.74% Mismatches: 1
Query Match: 82.73% Indels: 0
DB: 6 Gaps: 0

US-10-042-417A-91 (1-20) x US-08-196-363-9866 (1-275)
QY 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
Db 156 ATGCGCGTCTGTGGAGCAGAGCCCAAGAGCGCTGGCCCTCAGAAGAGCTCAACG 212

RESULT 5
US-08-196-363A-9866
; Sequence 9866, Application US/08196363A
; GENERAL INFORMATION:
; APPLICANT: Haseltine, William A.
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; APPLICANT: Dillon, Patrick J.
; APPLICANT: Li, Haodong
; APPLICANT: Adams, Mark D.
; APPLICANT: Bult, Carol
; APPLICANT: Blake, Judith A.
; APPLICANT: Clayton, Rebecca
; APPLICANT: Fitzgerald, Lisa M.
; APPLICANT: Fleischmann, Robert
; APPLICANT: Fraser, Claire M.
; APPLICANT: Fuldner, Rebecca A.
; APPLICANT: Gocayne, Jeannine D.
; APPLICANT: Kerlavage, Anthony R.
; APPLICANT: Kirkness, Ewen F.
; APPLICANT: Lee, Norman H.
; APPLICANT: Li, Li-Ing
; APPLICANT: Pellegrino, Susan M.
; APPLICANT: Sutton, Granger G.
; APPLICANT: White, Owen R.
; APPLICANT: Haseltine, William A.
; APPLICANT: Rosen, Steven M.
; APPLICANT: Dillon, Patrick J.
; APPLICANT: Li, Haodong
; APPLICANT: Adams, Mark D.
; APPLICANT: Bult, Carol
; APPLICANT: Blake, Judith A.
; APPLICANT: Clayton, Rebecca
; APPLICANT: Fitzgerald, Lisa M.
; APPLICANT: Fleischmann, Robert
; APPLICANT: Fraser, Claire M.
; APPLICANT: Fuldner, Rebecca A.
; APPLICANT: Gocayne, Jeannine D.
; APPLICANT: Kerlavage, Anthony R.
; APPLICANT: Kirkness, Ewen F.
; APPLICANT: Lee, Norman H.
; APPLICANT: Li, Li-Ing
; APPLICANT: Pellegrino, Susan M.
; APPLICANT: Sutton, Granger G.
; APPLICANT: White, Owen R.
; TITLE OF INVENTION: Human Genes, Sequences and Expression Products
; NUMBER OF SEQUENCES: 16820
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/196,363A
; FILING DATE: February 15, 1994
; CLASSIFICATION: 435
```

```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Kenley K. Hoover
; REGISTRATION NUMBER: 40,302
; REFERENCE/DOCKET NUMBER: PO-06
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 9866:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 275 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-196-363A-9866

Alignment Scores:
Pred. No.: 0.00244 Length: 275
Score: 91.00 Matches: 18
Percent Similarity: 94.74% Conservatve: 0
Best Local Similarity: 94.74% Mismatches: 1
Query Match: 82.73% Indels: 0
DB: 6 Gaps: 0

US-10-042-417A-91 (1-20) x US-08-196-363A-9866 (1-275)
QY 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
Db 157 AATGCCGGTCTGTGGAGCAGAGCCCAAGAGCGCTGGCCCTCAGAAGAGCTCAACG 213

RESULT 6
US-08-196-363-9866
; Sequence 9866, Application US/08196363D
; GENERAL INFORMATION:
; APPLICANT: Adams, Mark A.
; APPLICANT: Bult, Carol J.
; APPLICANT: Blake, Judith A.
; APPLICANT: Clayton, Rebecca
; APPLICANT: Fitzgerald, Lisa
; APPLICANT: Fleischmann, Robert D.
; APPLICANT: Fraser, Claire, M.
; APPLICANT: Fuldner, Rebecca A.
; APPLICANT: Gocayne, Jeannine D.
; APPLICANT: Kerlavage, Anthony R.
; APPLICANT: Kirkness, Ewen F.
; APPLICANT: Lee, Norman
; APPLICANT: Li, Li-Ing
; APPLICANT: Pellegrino, Susan
; APPLICANT: Sutton III, Granger G.
; APPLICANT: White, Owen R.
; APPLICANT: Haseltine, William A.
; APPLICANT: Rosen, Steven M.
; APPLICANT: Dillon, Patrick J.
; APPLICANT: Li, Haodong
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products
; FILE REFERENCE: PO-06
; CURRENT APPLICATION NUMBER: US/08/196,363D
; CURRENT FILING DATE: 1994-02-15
; EARLIER APPLICATION NUMBER:
; EARLIER FILING DATE:
; NUMBER OF SEQ ID NOS: 16820
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9866
; LENGTH: 275
; TYPE: DNA
; ORGANISM: Homo sapiens
US-08-196-363-9866

Alignment Scores:
Pred. No.: 0.00244 Length: 275
```

Score: 91.00 Matches: 18
Percent Similarity: 94.74% Conservative: 0
Best Local Similarity: 94.74% Mismatches: 1
Query Match: 82.73% Indels: 0
DB: 6 Gaps: 0

US-10-042-417A-91 (1-20) x US-08-196-363-9866 (1-275)

Qy 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
Db 157 AATCCGGTCTGTGGAGCAGACGCCCAAGAAGCCTGGCCTCAGAAGACGTCAAACG 213

RESULT 7

US-09-859-490-9866

; Sequence 9866, Application US/09859490
; GENERAL INFORMATION:
; APPLICANT: Adams, et al.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products
; FILE REFERENCE: PO-06C1
; CURRENT APPLICATION NUMBER: US/09/859,490
; CURRENT FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: 08/196,363
; PRIOR FILING DATE: 1994-02-15
; NUMBER OF SEQ ID NOS: 16820
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 9866
; LENGTH: 275
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-859-490-9866

Alignment Scores:

Pred. No.: 0.00244 Length: 275
Score: 91.00 Matches: 18
Percent Similarity: 94.74% Conservative: 0
Best Local Similarity: 94.74% Mismatches: 1
Query Match: 82.73% Indels: 0
DB: 35 Gaps: 0

US-10-042-417A-91 (1-20) x US-09-859-490-9866 (1-275)

Qy 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
Db 157 AATCCGGTCTGTGGAGCAGACGCCCAAGAAGCCTGGCCTCAGAAGACGTCAAACG 213

RESULT 8

US-09-912-293-115347

; Sequence 115347, Application US/09912293
; GENERAL INFORMATION:
; APPLICANT: Rosen, et. al.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 100
; FILE REFERENCE: PO-100
; CURRENT APPLICATION NUMBER: US/09/912,293
; CURRENT FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: 08/103,744
; PRIOR FILING DATE: 1993-08-09
; PRIOR APPLICATION NUMBER: 09/249,651
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: 08/104,507
; PRIOR FILING DATE: 1993-08-09
; PRIOR APPLICATION NUMBER: 08/196,363
; PRIOR FILING DATE: 1994-02-15
; PRIOR APPLICATION NUMBER: 09/859,490
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: 08/196,362
; PRIOR FILING DATE: 1994-02-15
; PRIOR APPLICATION NUMBER: 08/221,623
; PRIOR FILING DATE: 1994-03-31
; PRIOR APPLICATION NUMBER: 08/220,691
; PRIOR FILING DATE: 1994-03-31
; PRIOR APPLICATION NUMBER: 09/741,830
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 09/813,155

; PRIOR FILING DATE: 2001-03-21
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 244538
; SEQ ID NO 115347
; LENGTH: 275
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-912-293-115347

Alignment Scores:

Pred. NO.: 0.00244 Length: 275
Score: 91.00 Matches: 18
Percent Similarity: 94.74% Conservative: 0
Best Local Similarity: 94.74% Mismatches: 1
Query Match: 82.73% Indels: 0
DB: 38 Gaps: 0

US-10-042-417A-91 (1-20) x US-09-912-293-115347 (1-275)

Qy 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
Db 157 AATCCGGTCTGTGGAGCAGACGCCCAAGAAGCCTGGCCTCAGAAGACGTCAAACG 213

RESULT 9

US-60-253-653-5665

; Sequence 5665, Application US/60253653
; GENERAL INFORMATION:
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions isolated from ovine tissues
; FILE REFERENCE: 1057P
; CURRENT APPLICATION NUMBER: US/60/253,653
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 30124
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5665
; LENGTH: 364
; TYPE: DNA
; ORGANISM: Ovine
US-60-253-653-5665

Alignment Scores:

Pred. NO.: 0.00329 Length: 364
Score: 91.00 Matches: 18
Percent Similarity: 94.74% Conservative: 0
Best Local Similarity: 94.74% Mismatches: 1
Query Match: 82.73% Indels: 0
DB: 80 Gaps: 0

US-10-042-417A-91 (1-20) x US-60-253-653-5665 (1-364)

Qy 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
Db 157 AACCCCGGTTCAAGTGGAGCAGACGCCCAAGAAGCCTGGCCTCAGAAGCGTCAGACG 213

RESULT 10

US-09-397-022-3977

; Sequence 3977, Application US/09397022
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; APPLICANT: Holtzman, Douglas A.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
; FILE REFERENCE: MLN98-46pM
; CURRENT APPLICATION NUMBER: US/09/397,022
; CURRENT FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: 60/100,465
; PRIOR FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 60/106,443
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: 60/107,257
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: 60/126,906

```
; PRIOR FILING DATE: 1999-03-30
; PRIOR APPLICATION NUMBER: 60/132,099
; PRIOR FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 5775
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 3977
; LENGTH: 508
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(508)
; OTHER INFORMATION: n = A,T,C or G
US-09-397-022-3977

Alignment Scores:
Pred. No.: 0.00467 Length: 508
Score: 91.00 Matches: 18
Percent Similarity: 94.74% Conservative: 0
Best Local Similarity: 94.74% Mismatches: 1
Query Match: 82.73% Indels: 0
DB: 20 Gaps: 0

US-10-042-417A-91 (1-20) x US-09-397-022-3977 (1-508)
Qy 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
Db 313 AATGCCGGTCTGTGGAGCAGACGCCCAAGAGCGCTGGCCTCAGAGACGCTCAACG 369

RESULT 11
US-09-808-383-3977
; Sequence 3977, Application US/09808383
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; APPLICANT: Holtzman, Douglas A.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
; TITLE OF INVENTION: HUMAN FETAL SKIN LIBRARY
; FILE REFERENCE: 1600.1046-002
; CURRENT APPLICATION NUMBER: US/09/808,383
; CURRENT FILING DATE: 2001-03-13
; PRIOR APPLICATION NUMBER: US 09/397,022
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: US 60/100,465
; PRIOR FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: US 60/106,443
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: US 60/107,257
; PRIOR FILING DATE: 1998-11-15
; PRIOR APPLICATION NUMBER: US 60/126,906
; PRIOR FILING DATE: 1999-03-30
; PRIOR APPLICATION NUMBER: US 60/132,099
; PRIOR FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 5775
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 3977
; LENGTH: 508
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(508)
; OTHER INFORMATION: n = A,T,C or G
US-09-808-383-3977

Alignment Scores:
Pred. No.: 0.00467 Length: 508
Score: 91.00 Matches: 18
Percent Similarity: 94.74% Conservative: 0
Best Local Similarity: 94.74% Mismatches: 1
Query Match: 82.73% Indels: 0
DB: 34 Gaps: 0

US-10-042-417A-91 (1-20) x US-09-808-383-3977 (1-508)
```

```
Qy 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
Db 313 AATGCCGGTCTGTGGAGCAGACGCCCAAGAGCGCTGGCCTCAGAGACGCTCAACG 369

RESULT 12
US-09-724-676-18917
; Sequence 18917, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18917
; LENGTH: 562
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676-18917

Alignment Scores:
Pred. No.: 0.0052 Length: 562
Score: 91.00 Matches: 18
Percent Similarity: 94.74% Conservative: 0
Best Local Similarity: 94.74% Mismatches: 1
Query Match: 82.73% Indels: 0
DB: 32 Gaps: 0

US-10-042-417A-91 (1-20) x US-09-724-676-18917 (1-562)
Qy 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
Db 217 AATGCCGGTCTGTGGAGCAGACGCCCAAGAGCGCTGGCCTCAGAGACGCTCAACG 273

RESULT 13
US-09-724-676A-18917
; Sequence 18917, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18917
; LENGTH: 562
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676A-18917

Alignment Scores:
Pred. No.: 0.0052 Length: 562
Score: 91.00 Matches: 18
Percent Similarity: 94.74% Conservative: 0
Best Local Similarity: 94.74% Mismatches: 1
Query Match: 82.73% Indels: 0
DB: 32 Gaps: 0

US-10-042-417A-91 (1-20) x US-09-724-676A-18917 (1-562)
Qy 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
Db 217 AATGCCGGTCTGTGGAGCAGACGCCCAAGAGCGCTGGCCTCAGAGACGCTCAACG 273

RESULT 14
US-10-029-386-12066
; Sequence 12066, Application US/10029386
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
```


US-10-042-417A-91 (1-20) x US-09-534-843-34024 (1-580)
Qy 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
Db 80 AATCCCGTTCTGTGGAGCAGACGCCCAAGAAGCCTGGCCTCAGAGACGCTCAAACG 136

Search completed: August 28, 2003, 16:19:54
Job time : 3673 secs

APPLICANT: Hanzel, David K.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: ACOMICA-X-2
CURRENT APPLICATION NUMBER: US/10/029,386
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 34288
SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
SEQ ID NO 12066
LENGTH: 576
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO CHR12.3
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.1
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.6
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.3
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.2
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.6
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.6
OTHER INFORMATION: EST HUMAN HIT: U25937.1, EVALUE 0.00e+00
OTHER INFORMATION: SWISSPROT HIT: P46527, EVALUE 9.00e-17
OTHER INFORMATION: NT HIT: S76988.1, EVALUE 0.00e+00
US-10-029-386-12066

Alignment Scores:
Pred. No.: 0.00534 Length: 576
Score: 91.00 Matches: 18
Percent Similarity: 94.74% Conservative: 0
Best Local Similarity: 94.74% Mismatches: 1
Query Match: 82.73% Indels: 0
DB: 44 Gaps: 0

US-10-042-417A-91 (1-20) x US-10-029-386-12066 (1-576)

Qy 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
Db 208 AATCCCGTTCTGTGGAGCAGACGCCCAAGAAGCCTGGCCTCAGAGACGCTCAAACG 264

RESULT 15

US-09-534-843-34024
Sequence 34024, Application US/09534843
GENERAL INFORMATION:
APPLICANT: Seilhamer, Jeffrey J.
APPLICANT: Delegeane, Angelo M.
APPLICANT: Stuart, Susan G.
APPLICANT: Stuve, Laura L.
APPLICANT: Mullany, Sara J.
APPLICANT: Naughton, Rebecca E.
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING SIGNAL TRANSDUCTION MOLECULES
FILE REFERENCE: PD-1007 CIP
CURRENT APPLICATION NUMBER: US/09/534,843
CURRENT FILING DATE: 2000-03-24
Prior application data removed - refer to PALM or file wrapper
NUMBER OF SEQ ID NOS: 49783
SOFTWARE: PERL Program
SEQ ID NO 34024
LENGTH: 580
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID No: hu01156604
US-09-534-843-34024

Alignment Scores:
Pred. No.: 0.00538 Length: 580
Score: 91.00 Matches: 18
Percent Similarity: 94.74% Conservative: 0
Best Local Similarity: 94.74% Mismatches: 1
Query Match: 82.73% Indels: 0
DB: 23 Gaps: 0

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 28, 2003, 14:33:55 ; Search time 208 Seconds
(without alignments)
220.601 Million cell updates/sec

Title: US-10-042-417A-91

Perfect score: 110

Sequence: 1 NNAGSVEWTPKPKFLRRQT 20

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1533700 seqs, 1147125425 residues

Total number of hits satisfying chosen parameters: 3067400

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10042417/runat_19082003.141158_28034/app.query.fasta.1.199
-DB=PublishedApplications_NA -QFMT=fastap SUFFIX=rnpb -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT-1 -END=1 -MATRIX=Blissum62
-TRANS=human40.cd1 -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10042417.ecgn.1.1.238 @runat_19082003.141158_28034
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications_NA:**

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq**
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq**
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq**
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq**
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq**
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq**
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq**
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq**
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq**
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq**
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq**
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq**
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq**
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq**
15: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq**
16: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq**
17: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq**

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Score	Match Length	ID	Description
1	91	82.7	596	10 US-09-865-018-1
				Sequence 1, Appli

2	82.7	597	11	US-09-970-561-1	Sequence 1, Appli
3	82.7	1098	8	US-08-902-572-7	Sequence 7, Appli
4	82.7	1143	8	US-08-902-572-5	Sequence 5, Appli
5	82.7	1420	8	US-08-902-572-1	Sequence 1, Appli
6	79.1	284	10	US-09-867-701-3306	Sequence 3306, Ap
7	67.3	593	10	US-09-865-018-3	Sequence 3, Appli
8	67.3	309	9	US-09-728-446-1227	Sequence 1227, Ap
9	48.2	2940917	13	US-10-027-632-174763	Sequence 174763,
10	47.3	680	13	US-10-027-632-163306	Sequence 163306,
11	47.3	680	13	US-10-027-632-163307	Sequence 163307,
12	47.3	99014	10	US-09-880-107-3428	Sequence 3428, Ap
13	46.4	488	11	US-09-918-995-2359	Sequence 2359, Ap
14	46.4	2016	12	US-10-217-939-13	Sequence 13, Appli
15	45.9	1541	11	US-09-943-857-7	Sequence 7, Appli
16	45.5	524	13	US-10-027-632-324682	Sequence 324682,
17	45.5	571	9	US-09-864-761-16220	Sequence 16220, A
18	45.5	2364	10	US-09-712-363-32	Sequence 32, Appli
19	45.5	2370	14	US-10-205-823-298	Sequence 298, App
20	45.5	3027	14	US-10-310-002-8	Sequence 8, Appli
21	45.5	3408	10	US-09-870-759-133	Sequence 133, App
22	45.5	3408	12	US-09-751-708A-133	Sequence 133, App
23	45.5	3408	14	US-10-171-581-223	Sequence 223, App
24	45.5	3412	10	US-09-870-759-139	Sequence 139, App
25	45.5	3412	12	US-09-751-708A-139	Sequence 139, App
26	45.5	4926	14	US-10-310-002-5	Sequence 5, Appli
27	45.5	5848	14	US-10-310-002-44	Sequence 44, Appli
28	45.0	2017	14	US-10-102-806-137	Sequence 137, App
29	44.5	246	9	US-09-867-550-467	Sequence 467, App
30	44.5	277	9	US-09-923-876-281	Sequence 281, Ap
31	44.5	342	10	US-09-796-692-2697	Sequence 2697, Ap
32	44.5	342	14	US-10-040-862-2697	Sequence 2697, Ap
33	44.5	466	11	US-09-918-995-8784	Sequence 8784, Ap
34	44.5	533	13	US-10-027-632-246961	Sequence 246961,
35	44.5	1300	10	US-09-880-107-3731	Sequence 3731, Ap
36	44.5	1309	10	US-09-880-107-2235	Sequence 2235, Ap
37	44.5	713059	13	US-10-027-632-174581	Sequence 174581,
38	43.6	437	10	US-09-764-847-1342	Sequence 1342, Ap
39	43.6	437	14	US-10-092-154-1342	Sequence 1342, Ap
40	43.6	438	10	US-09-764-847-203	Sequence 203, App
41	43.6	438	14	US-10-092-154-203	Sequence 203, App
42	43.6	629	13	US-10-027-632-134770	Sequence 134770,
43	43.6	948	14	US-10-156-761-1604	Sequence 1604, Ap
44	43.6	1308	14	US-10-156-761-5472	Sequence 5472, Ap
45	43.6	5224	11	US-09-373-658-32	Sequence 32, Appli

ALIGNMENTS

RESULT 1

US-09-865-018-1
; Sequence 1, Application US/09865018
; Patent No: US20020110886A1
; GENERAL INFORMATION:
; APPLICANT: Massague, Joan
; Roberts, James M.
; Koff, Andrew
; Polyak, Kornelia

TITLE OF INVENTION: ISOLATED p27 PROTEIN AND METHOD FOR ITS
PRODUCTION AND USE

NUMBER OF SEQUENCES: 27

CORRESPONDENCE ADDRESS:

ADDRESSEE: FOLEY, HOAG & ELIOT LLP

STREET: One Post Office Square

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109-2170

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

```

; APPLICATION NUMBER: US/09/865.018
; FILING DATE: 24-May-2001
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US/08/854.039
; FILING DATE: 09-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MIV-079.04
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 596 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..594
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-865-018-1
Alignment Scores:
Pred. No.: 2.93e-05 Length: 596
Score: 91.00 Matches: 18
Percent Similarity: 94.74% Conservative: 0
Best Local Similarity: 94.74% Mismatches: 1
Query Match: 82.73% Indels: 0
DB: 10 Gaps: 0

US-10-042-417A-91 (1-20) x US-09-865-018-1 (1-596)
QY 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
Db 538 AATGCCGGTCTGTGGAGACAGCCCAAGAGCCTGGCCTCAGAGACGTCAAACG 594

RESULT 2
; Sequence 1, Application US/09970561
; Publication No. US20030023034A1
; GENERAL INFORMATION:
; APPLICANT: Nandabalan, Krishnan
; TITLE OF INVENTION: p27(Kip-1)-FKBP-12 Protein Complexes
; FILE REFERENCE: Cura-14 CON (15966-514 CON)
; CURRENT APPLICATION NUMBER: US/09/970,561
; CURRENT FILING DATE: 2001-10-03
; PRIOR APPLICATION NUMBER: PCT/US99/13659
; PRIOR FILING DATE: 1999-06-18
; PRIOR APPLICATION NUMBER: USSN 09/719,755
; PRIOR FILING DATE: 2000-12-15
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 1
; LENGTH: 597
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(594)
US-09-970-561-1
Alignment Scores:
Pred. No.: 2.93e-05 Length: 597
Score: 91.00 Matches: 18
Percent Similarity: 94.74% Conservative: 0
Best Local Similarity: 94.74% Mismatches: 1
Query Match: 82.73% Indels: 0
DB: 11 Gaps: 0

APPLICATION NUMBER: US/09/865.018
FILING DATE: 24-May-2001
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/08/854.039
FILING DATE: 09-MAY-1997
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: MIV-079.04
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-1000
TELEFAX: 617-832-7000
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 596 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..594
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-865-018-1
Alignment Scores:
Pred. No.: 2.93e-05 Length: 596
Score: 91.00 Matches: 18
Percent Similarity: 94.74% Conservative: 0
Best Local Similarity: 94.74% Mismatches: 1
Query Match: 82.73% Indels: 0
DB: 10 Gaps: 0

US-10-042-417A-91 (1-20) x US-09-865-018-1 (1-596)
QY 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
Db 538 AATGCCGGTCTGTGGAGACAGCCCAAGAGCCTGGCCTCAGAGACGTCAAACG 594

RESULT 2
; Sequence 1, Application US/09970561
; Publication No. US20030023034A1
; GENERAL INFORMATION:
; APPLICANT: Nandabalan, Krishnan
; TITLE OF INVENTION: p27(Kip-1)-FKBP-12 Protein Complexes
; FILE REFERENCE: Cura-14 CON (15966-514 CON)
; CURRENT APPLICATION NUMBER: US/09/970,561
; CURRENT FILING DATE: 2001-10-03
; PRIOR APPLICATION NUMBER: PCT/US99/13659
; PRIOR FILING DATE: 1999-06-18
; PRIOR APPLICATION NUMBER: USSN 09/719,755
; PRIOR FILING DATE: 2000-12-15
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 1
; LENGTH: 597
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(594)
US-09-970-561-1
Alignment Scores:
Pred. No.: 2.93e-05 Length: 597
Score: 91.00 Matches: 18
Percent Similarity: 94.74% Conservative: 0
Best Local Similarity: 94.74% Mismatches: 1
Query Match: 82.73% Indels: 0
DB: 11 Gaps: 0

US-10-042-417A-91 (1-20) x US-09-970-561-1 (1-597)
QY 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
Db 538 AATGCCGGTCTGTGGAGACAGCCCAAGAGCCTGGCCTCAGAGACGTCAAACG 594

RESULT 3
US-08-902-572-7
; Sequence 7, Application US/08902572
; Publication No. US20020068706A1
; GENERAL INFORMATION:
; APPLICANT: Gyuris, Jenö
; APPLICANT: Lamphere, Lou
; APPLICANT: Beach, David H.
; TITLE OF INVENTION: INHIBITORS OF CELL-CYCLE PROGRESSION AND
; TITLE OF INVENTION: RELATED THERETO
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/902,572
; FILING DATE: 29-JUL-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MIV-069.03
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1098 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1095
US-08-902-572-7
Alignment Scores:
Pred. No.: 5.27e-05 Length: 1098
Score: 91.00 Matches: 18
Percent Similarity: 94.74% Conservative: 0
Best Local Similarity: 94.74% Mismatches: 1
Query Match: 82.73% Indels: 0
DB: 8 Gaps: 0

US-10-042-417A-91 (1-20) x US-08-902-572-7 (1-1098)
QY 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
Db 1039 AATGCCGGTCTGTGGAGACAGCCCAAGAGCCTGGCCTCAGAGACGTCAAACG 1095

RESULT 4
US-08-902-572-5
; Sequence 5, Application US/08902572
; Publication No. US20020068706A1
; GENERAL INFORMATION:
; APPLICANT: Gyuris, Jenö
; APPLICANT: Lamphere, Lou
```

```
; APPLICANT: Beach, David H.
; TITLE OF INVENTION: INHIBITORS OF CELL-CYCLE PROGRESSION AND
; TITLE OF INVENTION: RELATED THERETO
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/902,572
; FILING DATE: 29-JUL-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MIV-069.03
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1143 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1140
; US-08-902-572-5

Alignment Scores:
Pred. No.: 5.48e-05 Length: 1143
Score: 91.00 Matches: 18
Percent Similarity: 94.74% Conservative: 0
Best Local Similarity: 94.74% Mismatches: 1
Query Match: 82.73% Indels: 0
DB: 8 Gaps: 0

US-10-042-417A-91 (1-20) x US-08-902-572-5 (1-1143)

Qy 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
Db 1084 AATGCCGGTCTGTGGAGCAGACGCCCAAGAGCTGGCTCAGAGACGTCAAACG 1140

RESULT 5
US-08-902-572-1
; Sequence 1, Application US/08902572
; Publication No. US20020068706A1
; GENERAL INFORMATION:
; APPLICANT: Gyuris, Jenő
; APPLICANT: Lamphere, Lou
; APPLICANT: Beach, David H.
; TITLE OF INVENTION: INHIBITORS OF CELL-CYCLE PROGRESSION AND
; TITLE OF INVENTION: RELATED THERETO
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

```
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/902,572
; FILING DATE: 29-JUL-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MIV-069.03
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1420 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 4..1176
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 4..24
; OTHER INFORMATION: /product= "poly-His_Tag"
; US-08-902-572-1

Alignment Scores:
Pred. No.: 6.75e-05 Length: 1420
Score: 91.00 Matches: 18
Percent Similarity: 94.74% Conservative: 0
Best Local Similarity: 94.74% Mismatches: 1
Query Match: 82.73% Indels: 0
DB: 8 Gaps: 0

US-10-042-417A-91 (1-20) x US-08-902-572-1 (1-1420)

Qy 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
Db 559 AATGCCGGTCTGTGGAGCAGACGCCCAAGAGCTGGCTCAGAGACGTCAAACG 615

RESULT 6
US-09-867-701-3306
; Sequence 3306, Application US/09867701
; Patent No. US2002013237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210x21.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3306
; LENGTH: 284
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(284)
; OTHER INFORMATION: n = A,T,C or G
; US-09-867-701-3306

Alignment Scores:
Pred. No.: 5.86e-05 Length: 284
Score: 87.00 Matches: 17
Percent Similarity: 89.47% Conservative: 0
```

Best Local Similarity: 89.47% Mismatches: 2
Query Match: 79.09% Indels: 0
DB: 10 Gaps: 0

US-10-042-417A-91 (1-20) x US-09-867-701-3306 (1-284)

Qy 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
Db 99 AATNCCGGTCTGTGGAGCAGACCCCAAGAGCGCTCGAAGACGCTCAACG 155

RESULT 7

US-09-865-018-3
; Sequence 3, Application US/09865018
; Patent No. US20020110886A1
; GENERAL INFORMATION:
; APPLICANT: Massague, Joan
; Roberts, James M.
; Koff, Andrew
; Polyak, Kornelia
; TITLE OF INVENTION: ISOLATED p27 PROTEIN AND METHOD FOR ITS
; PRODUCTION AND USE
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/865,018
; FILING DATE: 24-May-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/854,039
; FILING DATE: 09-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MIV-079.04
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 593 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..591
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-09-865-018-3

Alignment Scores:
Pred. No.: 0.0116 Length: 593
Score: 74.00 Matches: 14
Percent Similarity: 94.12% Conservative: 2
Best Local Similarity: 82.35% Mismatches: 1
Query Match: 67.27% Indels: 0
DB: 10 Gaps: 0

US-10-042-417A-91 (1-20) x US-09-865-018-3 (1-593)

Qy 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArg 18
Db 538 AACGCTGGCAGCTGTGGAGCAGACCCCAAGAGCGCTTCGACGCCAG 588

RESULT 8

US-09-728-446-1227
; Sequence 1227, Application US/09728446
; Patent No. US20020081668A1
; GENERAL INFORMATION:
; APPLICANT: Friedlich, Glenn
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. US20020081668A1el Murine Polynucleotide Sequences
; FILE OF INVENTION: and Mutant Cells and Mutant Animals Defined Thereby
; FILE REFERENCE: LEX-0101-USA
; CURRENT APPLICATION NUMBER: US/09/728,446
; CURRENT FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/168,270
; PRIOR FILING DATE: 1999-12-01
; NUMBER OF SEQ ID NOS: 1461
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1227
; LENGTH: 309
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(309)
; OTHER INFORMATION: n = A,T,C or G
US-09-728-446-1227

Alignment Scores:
Pred. No.: 0.0727 Length: 309
Score: 67.00 Matches: 13
Percent Similarity: 87.50% Conservative: 1
Best Local Similarity: 81.25% Mismatches: 2
Query Match: 60.91% Indels: 0
DB: 9 Gaps: 0

US-10-042-417A-91 (1-20) x US-09-728-446-1227 (1-309)

Qy 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArg 17
Db 96 AACGCTGGNACTGTGGAGCAGACCCCAAGAGCGCTTCGACGC 143

RESULT 9

US-10-027-632-174763
; Sequence 174763, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 174763
; LENGTH: 2940917
; TYPE: DNA
; ORGANISM: Human
; FEATURE:

```
; NAME/KEY: misc_feature
; LOCATION: (1)...(2940917)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-174763

Alignment Scores:
Pred. No.: 6,74e+04 Length: 2940917
Score: 53.00 Matches: 9
Percent Similarity: 69.23% Conservative: 0
Best Local Similarity: 69.23% Mismatches: 4
Query Match: 48.18% Indels: 0
DB: 13 Gaps: 0

US-10-042-417A-91 (1-20) x US-10-027-632-174763 (1-2940917)
QY 1 AsnAlaGlySerValGluTrpThrProLysLysPro 13
DB 2885273 AATAGGCTACTGCGCGGAATGACCCCTAAGAAGCCT 2885311

RESULT 10
US-10-027-632-163306/C
; Sequence 163306, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 2000-04-20
; PRIOR FILING DATE: 2000-04-20
; PRIOR FILING DATE: 2000-03-29
; PRIOR FILING DATE: 2000-03-29
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 1999-11-23
; PRIOR FILING DATE: 1999-11-23
; PRIOR FILING DATE: 1999-09-28
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 163306
; LENGTH: 680
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-163306

Alignment Scores:
Pred. No.: 30.5 Length: 680
Score: 52.00 Matches: 9
Percent Similarity: 73.68% Conservative: 5
Best Local Similarity: 47.37% Mismatches: 5
Query Match: 47.27% Indels: 0
DB: 13 Gaps: 0

US-10-042-417A-91 (1-20) x US-10-027-632-163306 (1-680)
QY 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
DB 312 AACTCACATAGCGTGCTCTGGACACCTGAGCTGCTGTGTTAAAGRTAAAGGAACC 256

RESULT 11
US-10-027-632-163307/C
; Sequence 163307, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 2000-04-20
; PRIOR FILING DATE: 2000-04-20
; PRIOR FILING DATE: 2000-03-29
; PRIOR FILING DATE: 2000-03-29
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 1999-11-23
; PRIOR FILING DATE: 1999-11-23
; PRIOR FILING DATE: 1999-09-28
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 163306
; LENGTH: 680
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-163306

Alignment Scores:
Pred. No.: 30.5 Length: 680
Score: 52.00 Matches: 9
Percent Similarity: 73.68% Conservative: 5
Best Local Similarity: 47.37% Mismatches: 5
Query Match: 47.27% Indels: 0
DB: 13 Gaps: 0

US-10-042-417A-91 (1-20) x US-10-027-632-163306 (1-680)
QY 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
DB 312 AACTCACATAGCGTGCTCTGGACACCTGAGCTGCTGTGTTAAAGRTAAAGGAACC 256

RESULT 12
US-09-880-107-3428
; Sequence 3428, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Marcel T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3428
; LENGTH: 99014
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 U78027
US-09-880-107-3428

Alignment Scores:
Pred. No.: 3,7e+03 Length: 99014
Score: 52.00 Matches: 9
Percent Similarity: 64.71% Conservative: 2
Best Local Similarity: 52.94% Mismatches: 6
Query Match: 47.27% Indels: 0
DB: 10 Gaps: 0
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 28, 2003, 13:38:04 ; Search time 65 Seconds
(without alignments)
135.810 Million cell updates/sec

Title: US-10-042-417A-91

Perfect score: 110

Sequence: 1 NNAGSVEWTPKPLRRRQT 20

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-O=/cgn2_1/USPTO.spool/US10042417/runat_19082003_141156_27948/app_query.fasta_1.199
-DB=Issued_Patents_NA -QPMF=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10042417 -CGN_1_1_56 @runat_19082003_141156_27948 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	91	82.7	596	4	US-08-854-039B-1
2	91	82.7	597	1	US-08-406-248-3
3	91	82.7	597	3	US-08-897-333A-1
4	91	82.7	597	3	US-08-240-906-5
5	91	82.7	597	4	US-08-794-002-1
6	91	82.7	597	4	US-09-457-568-25
7	91	82.7	597	4	US-09-457-646-25
8	91	82.7	597	4	US-09-516-065-25
9	91	82.7	1098	4	US-09-457-568-9
10	91	82.7	1098	4	US-09-457-646-9
11	91	82.7	1098	4	US-08-902-572-7
12	91	82.7	1098	4	US-09-516-065-9

13	91	82.7	1121	4	US-09-457-568-5	Sequence 5, Appli
14	91	82.7	1121	4	US-09-457-646-5	Sequence 5, Appli
15	91	82.7	1121	4	US-09-516-065-5	Sequence 5, Appli
16	91	82.7	1143	4	US-09-457-568-7	Sequence 7, Appli
17	91	82.7	1143	4	US-09-457-646-7	Sequence 7, Appli
18	91	82.7	1143	4	US-08-902-572-5	Sequence 5, Appli
19	91	82.7	1143	4	US-09-516-065-7	Sequence 7, Appli
20	91	82.7	1420	1	US-08-589-981-1	Sequence 1, Appli
21	91	82.7	1420	4	US-09-457-568-3	Sequence 3, Appli
22	91	82.7	1420	4	US-09-457-646-3	Sequence 3, Appli
23	91	82.7	1420	4	US-08-902-572-1	Sequence 1, Appli
24	91	82.7	1420	4	US-09-516-065-3	Sequence 3, Appli
25	74	67.3	593	4	US-08-854-039B-3	Sequence 3, Appli
26	74	67.3	594	3	US-09-240-906-7	Sequence 7, Appli
27	74	67.3	594	4	US-08-794-002-3	Sequence 3, Appli
28	51.5	46.8	4332	3	US-08-728-603-14	Sequence 14, Appl
29	51.5	46.8	32207	2	US-08-770-379-20	Sequence 20, Appl
30	51.5	46.8	32207	3	US-08-757-669A-20	Sequence 20, Appl
31	51.5	46.8	32207	4	US-09-230-371A-20	Sequence 20, Appl
32	51	46.4	476	1	US-07-947-120-9	Sequence 9, Appli
33	51	46.4	476	1	US-08-472-893A-9	Sequence 9, Appli
34	51	46.4	476	3	US-08-947-492-9	Sequence 9, Appli
35	50	45.5	5100	1	US-08-457-245-1	Sequence 1, Appli
36	50	45.5	4403765	3	US-09-103-840A-2	Sequence 2, Appli
37	50	45.5	4403765	3	US-09-103-840A-2	Sequence 2, Appli
38	50	45.5	4411529	3	US-09-103-840A-1	Sequence 1, Appli
39	50	45.5	4411529	3	US-09-103-840A-1	Sequence 1, Appli
40	48	43.6	863	2	US-08-673-190A-1	Sequence 1, Appli
41	48	43.6	2074	2	US-09-018-576-2	Sequence 2, Appli
42	48	43.6	2074	2	US-09-018-576-11	Sequence 11, Appl
43	48	43.6	2074	3	US-09-248-137-2	Sequence 2, Appli
44	48	43.6	2074	3	US-09-248-137-11	Sequence 11, Appl
45	48	43.6	2881	1	US-08-396-479B-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1

US-08-854-039B-1
; Sequence 1, Application US/08854039B
; Patent No. 6355774
; GENERAL INFORMATION:
; APPLICANT: Massague, Joan
; APPLICANT: Roberts, James M.
; APPLICANT: Koff, Andrew
; APPLICANT: Polyak, Kornelia
; TITLE OF INVENTION: ISOLATED p27 PROTEIN AND METHOD FOR ITS
; TITLE OF INVENTION: PRODUCTION AND USE
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/854,039B
; FILING DATE: 09-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MIV-079.04
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:

```
; LENGTH: 596 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..594
US-08-854-039B-1

Alignment Scores:
Pred. No.: 1.23e-05 Length: 596
Score: 91.00 Matches: 18
Percent Similarity: 94.74% Conservative: 0
Best Local Similarity: 94.74% Mismatches: 1
Query Match: 82.73% Indels: 0
DB: 4 Gaps: 0

US-10-042-417A-91 (1-20) x US-08-854-039B-1 (1-596)
Qy 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
Db 538 AATGCCGGTCTGTGGAGCAGACGCCCAAGAGCGCTGCCTCAGAGACGTCACACG 594

RESULT 2
US-08-406-248-3
; Sequence 3, Application US/08406248
; Patent No. 5736318
; GENERAL INFORMATION:
; APPLICANT: Munger, Karl
; APPLICANT: Jones, D. Leanne
; TITLE OF INVENTION: METHOD AND KIT FOR EVALUATING
; TITLE OF INVENTION: TRANSFORMED CELLS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ann-Louise Kerner, Ph.D., Lappin & Kusmer
; STREET: 200 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/406,248
; FILING DATE:
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: McDaniels, Patricia A.
; REGISTRATION NUMBER: 33,194
; REFERENCE/DOCKET NUMBER: HAZ-011
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-330-1300
; TELEFAX: 617-330-1311
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 597 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; DEVELOPMENTAL STAGE: Adult
; TISSUE TYPE: Kidney
; IMMEDIATE SOURCE:
; CLONE: p27kip1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..597
```

```
; PUBLICATION INFORMATION:
; AUTHORS: Polyak, K.
; AUTHORS: Lee, Mong-Hong
; AUTHORS: Erdjument-Bromage, H.
; AUTHORS: Koff, A.
; AUTHORS: Roberts, J. M.
; AUTHORS: Tempst, P.
; AUTHORS: Massague, J.
; TITLE: Cloning of p27kip1, a Cyclin-dependent Kinase
; TITLE: Inhibitor, and a Potential Mediator of
; TITLE: Extracellular Antimitogenic Signals
; JOURNAL: Cell
; VOLUME: 78
; PAGES: 56-66
; DATE: 1994
; RELEVANT RESIDUES IN SEQ ID NO: 3: FROM 1 TO 597
US-08-406-248-3

Alignment Scores:
Pred. No.: 1.23e-05 Length: 597
Score: 91.00 Matches: 18
Percent Similarity: 94.74% Conservative: 0
Best Local Similarity: 94.74% Mismatches: 1
Query Match: 82.73% Indels: 0
DB: 1 Gaps: 0

US-10-042-417A-91 (1-20) x US-08-406-248-3 (1-597)
Qy 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
Db 538 AATGCCGGTCTGTGGAGCAGACGCCCAAGAGCGCTGCCTCAGAGACGTCACACG 594

RESULT 3
US-08-897-333A-1
; Sequence 1, Application US/08897333A
; Patent No. 6177272
; GENERAL INFORMATION:
; APPLICANT: Nabel, Gary J.
; APPLICANT: Nabel, Elizabeth G.
; TITLE OF INVENTION: METHOD FOR TREATING VASCULAR PROLIFERATIVE DISEASES
; TITLE OF INVENTION: WITH P27 AND FUSIONS THEREOF
; FILE REFERENCE: 8642/4
; CURRENT APPLICATION NUMBER: US/08/897,333A
; CURRENT FILING DATE: 1998-07-21
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 597
; TYPE: DNA
; ORGANISM: Homo sapiens
US-08-897-333A-1

Alignment Scores:
Pred. No.: 1.23e-05 Length: 597
Score: 91.00 Matches: 18
Percent Similarity: 94.74% Conservative: 0
Best Local Similarity: 94.74% Mismatches: 1
Query Match: 82.73% Indels: 0
DB: 3 Gaps: 0

US-10-042-417A-91 (1-20) x US-08-897-333A-1 (1-597)
Qy 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
Db 538 AATGCCGGTCTGTGGAGCAGACGCCCAAGAGCGCTGCCTCAGAGACGTCACACG 594

RESULT 4
US-09-240-906-5
; Sequence 5, Application US/09240906
; Patent No. 6245965
; GENERAL INFORMATION:
; APPLICANT: ROUSSEL, MARTINE F.
; APPLICANT: SMEYNE, RICHARD
```

```
; APPLICANT: ZINDY, FREDERIQUE
; APPLICANT: CUNNINGHAM, JUSTINE
; TITLE OF INVENTION: CELLS THAT LACK PI3INK4D AND P27KIP1 ACTIVITY AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 1340-1-025
; CURRENT APPLICATION NUMBER: US/09/240,906
; CURRENT FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 597
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-240-906-5

Alignment Scores:
Pred. No.: 1.23e-05 Length: 597
Score: 91.00 Matches: 18
Percent Similarity: 94.74% Conservative: 0
Best Local Similarity: 94.74% Mismatches: 1
Query Match: 82.73% Indels: 0
DB: 3 Gaps: 0

US-10-042-417A-91 (1-20) x US-09-240-906-5 (1-597)

Qy 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
||||| 538 AATGCCGGTCTGTGGAGCAGACGCCCAAGAAGCTGGCCTCAGAGACGCTCAACG 594

RESULT 5
US-08-794-002-1
; Sequence 1, Application US/08794002
; Patent No. 6316208
; GENERAL INFORMATION:
; APPLICANT: Roberts, James M.
; APPLICANT: Porter, Peggy L.
; TITLE OF INVENTION: ISOLATED P27 PROTEIN AND METHODS FOR ITS
; TITLE OF INVENTION: PRODUCTION AND USE
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/794,002
; FILING DATE: 03-FEB-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MIV-079.03
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 597 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..594
US-08-794-002-1
```

```
Alignment Scores:
Pred. No.: 1.23e-05 Length: 597
Score: 91.00 Matches: 18
Percent Similarity: 94.74% Conservative: 0
Best Local Similarity: 94.74% Mismatches: 1
Query Match: 82.73% Indels: 0
DB: 4 Gaps: 0

US-10-042-417A-91 (1-20) x US-08-794-002-1 (1-597)

Qy 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
||||| 538 AATGCCGGTCTGTGGAGCAGACGCCCAAGAAGCTGGCCTCAGAGACGCTCAACG 594

RESULT 6
US-09-457-568-25
; Sequence 25, Application US/09457568
; Patent No. 6413943
; GENERAL INFORMATION:
; APPLICANT: McArthur, James G
; APPLICANT: Gyuris, Jen0
; APPLICANT: Finer, Mitchell H
; TITLE OF INVENTION: Methods and Reagents for Inhibiting Proliferation of
; TITLE OF INVENTION: Smooth Muscle Cells
; FILE REFERENCE: 106482.691
; CURRENT APPLICATION NUMBER: US/09/457,568
; CURRENT FILING DATE: 1999-12-09
; EARLIER APPLICATION NUMBER: 60/122,974
; EARLIER FILING DATE: 1999-03-01
; EARLIER APPLICATION NUMBER: 60/163,682
; EARLIER FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 25
; LENGTH: 597
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-457-568-25

Alignment Scores:
Pred. No.: 1.23e-05 Length: 597
Score: 91.00 Matches: 18
Percent Similarity: 94.74% Conservative: 0
Best Local Similarity: 94.74% Mismatches: 1
Query Match: 82.73% Indels: 0
DB: 4 Gaps: 0

US-10-042-417A-91 (1-20) x US-09-457-568-25 (1-597)

Qy 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
||||| 538 AATGCCGGTCTGTGGAGCAGACGCCCAAGAAGCTGGCCTCAGAGACGCTCAACG 594

RESULT 7
US-09-457-646-25
; Sequence 25, Application US/09457646
; Patent No. 6420345
; GENERAL INFORMATION:
; APPLICANT: Patel, Salil D
; APPLICANT: McArthur, James G
; APPLICANT: Gyuris, Jen0
; TITLE OF INVENTION: Methods and Reagents for Inhibiting Proliferation of
; TITLE OF INVENTION: Smooth Muscle Cells
; FILE REFERENCE: 106482.287
; CURRENT APPLICATION NUMBER: US/09/457,646
; CURRENT FILING DATE: 1999-12-09
; EARLIER APPLICATION NUMBER: 60/122,974
; EARLIER FILING DATE: 1999-03-01
; EARLIER APPLICATION NUMBER: 60/163,682
; EARLIER FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
```

```
; SEQ ID NO 25
; LENGTH: 597
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-457-646-25

Alignment Scores:
Pred. No.: 1.23e-05 Length: 597
Score: 91.00 Matches: 18
Percent Similarity: 94.74% Conservative: 0
Best Local Similarity: 94.74% Mismatches: 1
Query Match: 82.73% Indels: 0
DB: 4 Gaps: 0

US-10-042-417A-91 (1-20) x US-09-457-646-25 (1-597)
QY 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
Db 538 AATGCCGGTCTCTGGAGCAGAGCCGCCAAGAAGCCTGGCCCTCAGAAGACGTCAAACG 594

RESULT 8
US-09-516-065-25
; Sequence 25, Application US/09516065
; Patent No. 6521602
; GENERAL INFORMATION:
; APPLICANT: Patel, Salil D
; APPLICANT: McArthur, James G
; APPLICANT: Gyuris, Jen0
; APPLICANT: Mendez, Michel
; APPLICANT: Finer, Mitchell
; TITLE OF INVENTION: Anti-Neoplastic Compositions and Uses Thereof
; FILE REFERENCE: Cell 406; 106482.692
; CURRENT APPLICATION NUMBER: US/09/516,065
; CURRENT FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/128.515
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/128,271
; PRIOR FILING DATE: 1999-04-08
; PRIOR APPLICATION NUMBER: 60/122,974
; PRIOR FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 25
; LENGTH: 597
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-516-065-25

Alignment Scores:
Pred. No.: 1.23e-05 Length: 597
Score: 91.00 Matches: 18
Percent Similarity: 94.74% Conservative: 0
Best Local Similarity: 94.74% Mismatches: 1
Query Match: 82.73% Indels: 0
DB: 4 Gaps: 0

US-10-042-417A-91 (1-20) x US-09-516-065-25 (1-597)
QY 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
Db 538 AATGCCGGTCTCTGGAGCAGAGCCGCCAAGAAGCCTGGCCCTCAGAAGACGTCAAACG 594

RESULT 9
US-09-457-568-9
; Sequence 9, Application US/09457568
; Patent No. 6413943
; GENERAL INFORMATION:
; APPLICANT: McArthur, James G
; APPLICANT: Gyuris, Jen0
; APPLICANT: Finer, Mitchell H
; TITLE OF INVENTION: Methods and Reagents for Inhibiting Proliferation of
; FILE REFERENCE: 106482.691

; SEQ ID NO 25
; LENGTH: 597
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-457-568-9

Alignment Scores:
Pred. No.: 1.23e-05 Length: 597
Score: 91.00 Matches: 18
Percent Similarity: 94.74% Conservative: 0
Best Local Similarity: 94.74% Mismatches: 1
Query Match: 82.73% Indels: 0
DB: 4 Gaps: 0

US-10-042-417A-91 (1-20) x US-09-457-568-9 (1-1098)
QY 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
Db 1039 AATGCCGGTCTCTGGAGCAGAGCCGCCAAGAAGCCTGGCCCTCAGAAGACGTCAAACG 1095

RESULT 10
US-09-457-646-9
; Sequence 9, Application US/09457646
; Patent No. 6420345
; GENERAL INFORMATION:
; APPLICANT: Patel, Salil D
; APPLICANT: McArthur, James G
; APPLICANT: Gyuris, Jen0
; TITLE OF INVENTION: Methods and Reagents for Inhibiting Proliferation of
; FILE REFERENCE: 106482.287
; CURRENT APPLICATION NUMBER: US/09/457,646
; CURRENT FILING DATE: 1999-12-09
; EARLIER APPLICATION NUMBER: 60/122,974
; EARLIER FILING DATE: 1999-03-01
; EARLIER APPLICATION NUMBER: 60/163,682
; EARLIER FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 1098
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-457-646-9

Alignment Scores:
Pred. No.: 2.49e-05 Length: 1098
Score: 91.00 Matches: 18
Percent Similarity: 94.74% Conservative: 0
Best Local Similarity: 94.74% Mismatches: 1
Query Match: 82.73% Indels: 0
DB: 4 Gaps: 0

US-10-042-417A-91 (1-20) x US-09-457-646-9 (1-1098)
QY 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
Db 1039 AATGCCGGTCTCTGGAGCAGAGCCGCCAAGAAGCCTGGCCCTCAGAAGACGTCAAACG 1095

RESULT 11
US-08-902-572-7
; Sequence 7, Application US/08902572
; Patent No. 6495526
; GENERAL INFORMATION:
; APPLICANT: Gyuris, Jen0
```

APPLICANT: Lamphere, Lou
 APPLICANT: Beach, David H.
 TITLE OF INVENTION: INHIBITORS OF CELL-CYCLE PROGRESSION AND
 TITLE OF INVENTION: RELATED THERETO
 NUMBER OF SEQUENCES: 34
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: FOLEY, HOAG & ELIOT LLP
 STREET: One Post Office Square
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02109-2170
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION NUMBER: US/08/902,572
 FILING DATE: 29-JUL-1997
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: Vincent, Matthew P.
 REGISTRATION NUMBER: 36,709
 REFERENCE/DOCKET NUMBER: MIV-069.03
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-832-1000
 TELEFAX: 617-832-7000
 INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1098 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: both
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..1095
 US-08-902-572-7

Alignment Scores:
 Pred. No.: 2.49e-05 Length: 1098
 Score: 91.00 Matches: 18
 Percent Similarity: 94.74% Conservative: 0
 Best Local Similarity: 94.74% Mismatches: 1
 Query Match: 82.73% Indels: 0
 DB: 0 Gaps: 0

US-10-042-417A-91 (1-20) x US-08-902-572-7 (1-1098)

QY 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
 Db 1039 AATGCCGGTCTCTGTGGAGCAGACGCCCAAGAGCCTGGCCTCAGAGACGTCAAACG 1095

RESULT 12

US-09-516-065-9
 ; Sequence 9, Application US/09516065
 ; Patent No. 6521602
 ; GENERAL INFORMATION:
 ; APPLICANT: Patel, Salil D
 ; APPLICANT: McArthur, James G
 ; APPLICANT: Gyuris, Jenio
 ; APPLICANT: Mendez, Michel
 ; APPLICANT: Finer, Mitchel
 ; TITLE OF INVENTION: Anti-Neoplastic Compositions and Uses Thereof
 ; FILE REFERENCE: Cell 406; 106482.692
 ; CURRENT APPLICATION NUMBER: US/09/516,065
 ; CURRENT FILING DATE: 2000-03-01
 ; PRIOR FILING DATE: 1999-04-09
 ; PRIOR APPLICATION NUMBER: 60/128,515
 ; PRIOR FILING DATE: 1999-04-09
 ; PRIOR APPLICATION NUMBER: 60/128,271
 ; PRIOR FILING DATE: 1999-04-08
 ; PRIOR APPLICATION NUMBER: 60/122,974

; PRIOR FILING DATE: 1999-03-01
 ; NUMBER OF SEQ ID NOS: 32
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 9
 ; LENGTH: 1098
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-516-065-9

Alignment Scores:
 Pred. No.: 2.49e-05 Length: 1098
 Score: 91.00 Matches: 18
 Percent Similarity: 94.74% Conservative: 0
 Best Local Similarity: 94.74% Mismatches: 1
 Query Match: 82.73% Indels: 0
 DB: 0 Gaps: 0

US-10-042-417A-91 (1-20) x US-09-516-065-9 (1-1098)

QY 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
 Db 1039 AATGCCGGTCTCTGTGGAGCAGACGCCCAAGAGCCTGGCCTCAGAGACGTCAAACG 1095

RESULT 13

US-09-457-568-5
 ; Sequence 5, Application US/09457568
 ; Patent No. 6413943
 ; GENERAL INFORMATION:
 ; APPLICANT: McArthur, James G
 ; APPLICANT: Gyuris, Jenio
 ; APPLICANT: Finer, Mitchell H
 ; TITLE OF INVENTION: Methods and Reagents for Inhibiting Proliferation of
 ; TITLE OF INVENTION: Smooth Muscle Cells
 ; FILE REFERENCE: 106482.691
 ; CURRENT APPLICATION NUMBER: US/09/457,568
 ; CURRENT FILING DATE: 1999-12-09
 ; EARLIER APPLICATION NUMBER: 60/122,974
 ; EARLIER FILING DATE: 1999-03-01
 ; EARLIER APPLICATION NUMBER: 60/163,682
 ; EARLIER FILING DATE: 1999-11-05
 ; NUMBER OF SEQ ID NOS: 28
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 5
 ; LENGTH: 1121
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-457-568-5

Alignment Scores:

Pred. No.: 2.55e-05 Length: 1121
 Score: 91.00 Matches: 18
 Percent Similarity: 94.74% Conservative: 0
 Best Local Similarity: 94.74% Mismatches: 1
 Query Match: 82.73% Indels: 0
 DB: 0 Gaps: 0

US-10-042-417A-91 (1-20) x US-09-457-568-5 (1-1121)

QY 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
 Db 583 AATGCCGGTCTCTGTGGAGCAGACGCCCAAGAGCCTGGCCTCAGAGACGTCAAACG 639

RESULT 14

US-09-457-646-5
 ; Sequence 5, Application US/09457646
 ; Patent No. 6420345
 ; GENERAL INFORMATION:
 ; APPLICANT: Patel, Salil D
 ; APPLICANT: McArthur, James G
 ; APPLICANT: Gyuris, Jenio
 ; TITLE OF INVENTION: Methods and Reagents for Inhibiting Proliferation of
 ; TITLE OF INVENTION: Smooth Muscle Cells
 ; FILE REFERENCE: 106482.287

; CURRENT APPLICATION NUMBER: US/09/457,646
 ; CURRENT FILING DATE: 1999-12-09
 ; EARLIER APPLICATION NUMBER: 60/122,974
 ; EARLIER FILING DATE: 1999-03-01
 ; EARLIER APPLICATION NUMBER: 60/163,682
 ; EARLIER FILING DATE: 1999-11-05
 ; NUMBER OF SEQ ID NOS: 32
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 5
 ; LENGTH: 1121
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-457-646-5

Alignment Scores:
 Pred. No.: 2.55e-05 Length: 1121
 Score: 91.00 Matches: 18
 Percent Similarity: 94.74% Conservative: 0
 Best Local Similarity: 94.74% Mismatches: 1
 Query Match: 82.73% Indels: 0
 DB: 4 Gaps: 0

US-10-042-417A-91 (1-20) x US-09-457-646-5 (1-1121)

QY 2 AsnAlaGlySerValGluTrpThrProLysProGlyLeuArgArgGlnThr 20
 Db 583 AATGCCGGTCTGTGGAGCAGACGCCCAAGAGCCTGGCCTCAGAGACGTCACACG 639

RESULT 15

US-09-516-065-5
 ; Sequence 5, Application US/09516065
 ; Patent No. 6521602
 ; GENERAL INFORMATION:
 ; APPLICANT: Patel, Salil D
 ; APPLICANT: McArthur, James G
 ; APPLICANT: Gyuris, Jeno
 ; APPLICANT: Mendez, Michal
 ; APPLICANT: Finer, Michel
 ; TITLE OF INVENTION: Anti-Neoplastic Compositions and Uses Thereof
 ; FILE REFERENCE: Cell 406; 106482.692
 ; CURRENT APPLICATION NUMBER: US/09/516,065
 ; CURRENT FILING DATE: 2000-03-01
 ; PRIOR APPLICATION NUMBER: 60/128,515
 ; PRIOR FILING DATE: 1999-04-09
 ; PRIOR APPLICATION NUMBER: 60/128,271
 ; PRIOR FILING DATE: 1999-04-08
 ; PRIOR APPLICATION NUMBER: 60/122,974
 ; PRIOR FILING DATE: 1999-03-01
 ; NUMBER OF SEQ ID NOS: 32
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 5
 ; LENGTH: 1121
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-516-065-5

Alignment Scores:
 Pred. No.: 2.55e-05 Length: 1121
 Score: 91.00 Matches: 18
 Percent Similarity: 94.74% Conservative: 0
 Best Local Similarity: 94.74% Mismatches: 1
 Query Match: 82.73% Indels: 0
 DB: 4 Gaps: 0

US-10-042-417A-91 (1-20) x US-09-516-065-5 (1-1121)

QY 2 AsnAlaGlySerValGluTrpThrProLysProGlyLeuArgArgGlnThr 20
 Db 583 AATGCCGGTCTGTGGAGCAGACGCCCAAGAGCCTGGCCTCAGAGACGTCACACG 639

Search completed: August 28, 2003, 15:18:42
 Job time : 81 secs

GenCore version 5.1.6.
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 28, 2003, 13:28:34 ; Search time 284 Seconds
(without alignments)
190.101 Million cell updates/sec

Title: US-10-042-417A-91
Perfect score: 110
Sequence: 1 NNAGSVEWTPKPKPLRRRQT 20

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-O=/cgn2.1/USPTO.spool/US10042417/runat.19082003.141154.27878/app.query.fasta.1.199
-DB=N_Geneseq.19Jun03 -QPM=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blom62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10042417 -CGN.1.1.312 -runat.19082003.141154.27878 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq.19Jun03.*
1: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed.

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	91	82.7	596	17	Human p27 Kipl CDN
2	91	82.7	597	19	CDNA encoding prot
3	91	82.7	597	19	Nucleotide sequenc
4	91	82.7	597	20	CKI/KIP protein p2
5	91	82.7	597	21	Human p27 protein
6	91	82.7	597	21	Human cyclin depen
7	91	82.7	597	21	Human cyclin depen
8	91	82.7	597	21	Human mutant cycli
9	91	82.7	597	21	Human p27(Kipl) ki
10	91	82.7	597	22	Nucleotide sequenc
11	91	82.7	597	22	Human p27 protein
12	91	82.7	597	24	Hypoxia-regulated
13	91	82.7	597	24	p27-Kipl CDNA. Ho
14	91	82.7	1098	18	CDK inhibitory fus
15	91	82.7	1098	20	Human p16p27 fusio
16	91	82.7	1098	21	Human W6 protein c
17	91	82.7	1098	21	DNA encoding p16-p
18	91	82.7	1098	21	DNA encoding angio
19	91	82.7	1121	20	Human p16p27 fusio
20	91	82.7	1121	21	Human W4 protein c
21	91	82.7	1121	21	DNA encoding p27-p
22	91	82.7	1121	21	DNA encoding angio
23	91	82.7	1143	18	CDK inhibitory fus
24	91	82.7	1143	20	Human p16(GS)p27 f
25	91	82.7	1143	21	Human W5 protein c
26	91	82.7	1143	21	DNA encoding p16-p
27	91	82.7	1143	21	DNA encoding angio
28	91	82.7	1420	18	CDK inhibitory fus
29	91	82.7	1420	20	Human p27-p16 fusi
30	91	82.7	1420	21	Human W3 protein c
31	91	82.7	1420	21	DNA encoding p27-p
32	91	82.7	1420	21	DNA encoding angio
33	87	79.1	284	24	Human ovarian canc
34	82	74.5	1958	24	Porcine p27Kipl nu
35	75	68.2	60	24	Human spliced tran
36	74	67.3	199	22	Murine transcripti
37	74	67.3	593	17	Mouse p27 Kipl CDN
38	74	67.3	594	19	CDNA encoding prot
39	74	67.3	594	22	Nucleotide sequenc
40	71	64.5	582	20	DNA encoding p27 p
41	67	60.9	309	24	Mouse ES cell rela
42	53.5	48.6	429	23	Human prostate exp
43	52	47.3	99014	24	Gene #3429 used to
44	51.5	46.8	4332	19	Kaposi's sarcoma a
45	51.5	46.8	32207	20	XSHV LUR DNA (nucl

ALIGNMENTS

RESULT 1
AAT16336
ID AAT16336 standard; cDNA; 596 BP.
AC AAT16336;
XX
XX
DT 16-JUL-1996 (first entry)
DE Human p27 Kipl CDNA clone.
XX
XX p27 protein; Kipl; cyclin E; Cdk2; cell proliferation; ulcer;
KW cancer; hyperplasia; diagnosis; therapy; ss.
XX
XX Homo sapiens.
XX
XX W09602140-A1.
XX

PI Muenger K, Jones DL;
 XX WPI; 1998-239202/21.
 DR P-PSDB; AAW46888.
 XX
 XX Evaluation of proliferative state of cells transformed with human
 PT papilloma virus - by determining cyclin-dependent kinase activity
 PT induced by E7 onco-protein
 XX
 XX Disclosure; Columns 15-18; 14pp; English.
 XX
 CC The present sequence encodes a p27Kip1 protein, which is part of a
 CC family of small cyclin-dependent kinase inhibitors. The proliferative
 CC state of a cell transformed with Human papillomavirus (HPV) can be
 CC evaluated in the following manner. Cyclin/cyclin-dependent kinase
 CC complexes containing protein p27Kip1 are isolated from the transformed
 CC cell, and the HPV E7 oncoprotein (AAW46886) added to the isolated
 CC protein. Cyclin/cyclin-dependent kinase complexes are isolated from an
 CC untransformed cell that is substantially homogenic with the transformed
 CC cell, and the HPV E7 oncoprotein added. The kinase activities of the 2
 CC samples are measured, where a proliferating transformed cell has a
 CC greater kinase activity than the untransformed cell. The method is
 CC used for determining the extent of interaction and/or inactivation
 CC between a cyclin/cyclin-dependent kinase inhibitor and the HPV E7
 CC oncoprotein and thus evaluating the proliferative state of a transformed
 CC cell.
 CC (Updated on 25-MAR-2003 to correct PI field.)
 XX
 XX Sequence 597 BP; 161 A; 164 C; 185 G; 87 T; 0 other;
 SQ

Alignment Scores:
 Pred. NO.: 0.000172 Length: 597
 Score: 91.00 Matches: 18
 Percent Similarity: 94.74% Conservative: 0
 Best Local Similarity: 94.74% Mismatches: 1
 Query Match: 82.73% Indels: 0
 DB: 19 Gaps: 0

US-10-042-417A-91 (1-20) x AAV16719 (1-597)
 Qy 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
 AAX21817
 Db 538 AATCCCGTCTGTGGAGCAGACGCCCAAGAAGCTGGCTCAGAGACGTCAAACG 594
 |||||||
 AAX21817 standard; cDNA; 597 BP.
 AC AAX21817;
 XX
 XX 14-MAY-1999 (first entry)
 DT
 DE CKI/KIP protein p27 coding sequence.
 XX
 KW CKI/KIP protein; p27 protein; cyclin kinase inhibitor; cancer;
 KW hyperproliferative disorder; ds.
 XX
 OS Homo sapiens.
 XX
 PN W09904238-A2.
 XX
 XX 28-JAN-1999.
 PD
 XX 14-JUL-1998; 98WO-US14566.
 XX
 XX 15-JUL-1997; 97US-0893276.
 PR
 XX (DEAC-) DEACONESS HOSPITAL.
 PA (MITO-) MITOTIX INC.
 XX
 XX Draetta G, Loda M, Pagano M, Rolfe M;
 XX WPI; 1999-132426/11.
 DR

DR P-PSDB; AAY00768.
 XX
 XX Methods for diagnosis and prognosis of hyperproliferative disorders
 PT - by determining the level of cyclin kinase inhibitor protein(s),
 PT particularly p27
 XX
 XX Claim 19; Page 35-37; 53pp; English.
 XX
 CC This sequence encodes the cyclin kinase inhibitor (CKI) protein p27. The
 CC invention relates to a method for diagnosing a hyperproliferative
 CC disorder, associated with the destabilisation of a CKI protein in cells
 CC of a patient, comprises: (i) ascertaining the CKI protein level in a
 CC sample of patient cells; and (ii) diagnosing the presence or absence of a
 CC hyperproliferative disorder by utilising the ascertained CKI protein
 CC level, where a reduced CKI protein level, relative to a normal control
 CC cell sample, correlates with the presence of a hyperproliferative
 CC disorder. The methods are useful for diagnosing disorders associated with
 CC hyperproliferation, evaluating their aggressiveness and/or rate of
 CC recurrence and as prognosis for evaluating a cancer patient's risk of
 CC death. From the observations, treatment can be applied on the basis of
 CC the patient's risk of death and/or recurrence of the cancer. The
 CC diagnostic methods may also be employed as follow-up to treatment,
 CC e.g. quantitation of the level of p27 protein may be indicative of the
 CC effectiveness of current or previously employed cancer therapies as well
 CC as the effect of these therapies upon patient prognosis. The methods and
 CC reagents allow the detection of loss of p27 protein from a cell in order
 CC to diagnose and phenotype proliferative disorders arising from
 CC tumorigenic transformation of cells, or other hyperplastic or neoplastic
 CC transformation processes as well as differentiative disorders such as
 CC degeneration of tissue e.g. neurodegeneration.
 XX
 XX Sequence 597 BP; 161 A; 164 C; 185 G; 87 T; 0 other;
 SQ

Alignment Scores:
 Pred. NO.: 0.000172 Length: 597
 Score: 91.00 Matches: 18
 Percent Similarity: 94.74% Conservative: 0
 Best Local Similarity: 94.74% Mismatches: 1
 Query Match: 82.73% Indels: 0
 DB: 20 Gaps: 0

US-10-042-417A-91 (1-20) x AAX21817 (1-597)
 Qy 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
 |||||||
 Db 538 AATCCCGTCTGTGGAGCAGACGCCCAAGAAGCTGGCTCAGAGACGTCARACG 594
 |||||||
 RESULT 5
 AAA90920
 ID AAA90920 standard; DNA; 597 BP.
 XX
 XX AAA90920;
 AC
 XX
 XX 15-JAN-2001 (first entry)
 DT
 XX Human p27 protein coding sequence.
 DE
 XX
 KW Human; chimeric cyclin dependent kinase inhibitor; CDKi; therapy;
 KW adenovirus E4 protein; neoplasia; p27 protein; ss.
 XX
 OS Homo sapiens.
 XX
 XX W0200052184-A1.
 PN
 XX 08-SEP-2000.
 PD
 XX 01-MAR-2000; 2000WO-US05350.
 XX
 XX 01-MAR-1999; 99US-0122974.
 PR 08-APR-1999; 99US-0128271.
 PR 09-APR-1999; 99US-0128515.
 XX
 XX (CELL-) CELL GENESYS INC.
 PA

PA (MITO-) MITOTIX INC.
 XX
 PI Patel S, McArthur J, Gyuris J, Mendez MJ, Finer M;
 XX
 DR WPI; 2000-587315/55.
 DR P-PSDB; AAY97523.
 XX
 PT Protein and nucleic acid compositions for preventing and treating
 PT neoplasias (particularly cancer), comprises a novel chimeric cyclin
 PT dependent kinase inhibitor and adenovirus E4 protein
 XX
 PS Example 1; Page 122; 126pp; English.
 XX
 CC This sequence encodes the human p27 protein.
 CC The invention relates to a protein composition comprising a novel
 CC purified chimeric cyclin dependent kinase inhibitor (CDKi) and a
 CC purified adenovirus E4 protein. The compositions comprising the protein,
 CC or the DNA encoding it are useful for treating neoplasias in animals. The
 CC compositions also find use in assays to eliminate a specific
 CC sub-population of cultured cells, to determine the susceptibility of
 CC neoplastic cells to treatment with the compositions and also in assays to
 CC synchronise cell growth in cultured cells.
 XX
 SQ Sequence 597 BP; 161 A; 164 C; 185 G; 87 T; 0 other;

Alignment Scores:
 Pred. No.: 0.000172 Length: 597
 Score: 91.00 Matches: 18
 Percent Similarity: 94.74% Conservative: 0
 Best Local Similarity: 94.74% Mismatches: 1
 Query Match: 82.73% Indels: 0
 DB: 21 Gaps: 0

US-10-042-417A-91 (1-20) x AAA90920 (1-597)

Qy 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
 Db 538 AATGCCGGTCTGTGGAGCAGACGCCCAAGAGCGCTGCCTCAGAACGCTCAACG 594

RESULT 6
 AAA50499
 ID AAA50499 standard; cDNA; 597 BP.
 XX
 AC AAA50499;
 XX
 DT 05-DEC-2000 (first entry)
 XX
 DE Human cyclin dependent kinase inhibitor p27 DNA.
 XX
 KW Cyclin dependent kinase inhibitor; CDKi; CIP; KIP; human; p27;
 KW smooth muscle cell; restenosis; vasotropic; antiproliferative;
 KW gene therapy; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200052159-A1.
 XX
 PD 08-SEP-2000.
 XX
 PF 28-FEB-2000; 2000WO-US04971.
 XX
 PR 01-MAR-1999; 99US-0122974.
 PR 05-NOV-1999; 99US-0163682.
 PR 09-DEC-1999; 99US-0457568.
 XX
 PA (CELL-) CELL GENESYS INC.
 PA (MITO-) MITOTIX INC.
 XX
 PI McArthur J, Gyuris J, Finer M;
 XX
 DR WPI; 2000-594183/56.
 DR P-PSDB; AAY96052.
 XX

PT Novel recombinant lentivirus for inhibiting proliferation of smooth
 PT muscle cells in e.g. restenosis, is replication deficient and comprises
 PT a transgene encoding a cyclin dependent kinase inhibitor
 XX

XX Example 1; Page 119; 126pp; English.

XX The present sequence is that of DNA encoding human p27 (see
 CC AAY96052), a cyclin dependent kinase inhibitor (CDKi) that inhibits
 CC smooth muscle cell proliferation. A claimed method for inhibiting
 CC smooth muscle cell hyperproliferation involves transducing smooth
 CC muscle cells with a replication-deficient recombinant adenovirus
 CC that lacks a functional E1 region and a functional E4 region, and
 CC comprises a transgene encoding a CDKi. The CDKi is selected from
 CC an INK4 family protein such as human p16, a Cip/Kip family protein
 CC such as p27, active fragments of these, or fusion proteins
 CC comprising active fragments of) an INK4 family protein and a
 CC Cip/Kip family protein (see AAY96046 and AAY96049). The method is used
 CC to inhibit mammalian smooth muscle cell hyperproliferation, induced
 CC by injury caused by angioplasty, stent placement or vein
 CC engraftment. It is useful for treating vascular pathologies e.g.,
 CC restenosis. Also claimed are recombinant lentiviruses encoding
 CC CDKis.
 XX

SQ Sequence 597 BP; 161 A; 164 C; 185 G; 87 T; 0 other;

Alignment Scores:
 Pred. No.: 0.000172 Length: 597
 Score: 91.00 Matches: 18
 Percent Similarity: 94.74% Conservative: 0
 Best Local Similarity: 94.74% Mismatches: 1
 Query Match: 82.73% Indels: 0
 DB: 21 Gaps: 0

US-10-042-417A-91 (1-20) x AAA50499 (1-597)

Qy 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
 Db 538 AATGCCGGTCTGTGGAGCAGACGCCCAAGAGCGCTGCCTCAGAACGCTCAACG 594

RESULT 7
 AAA50519
 ID AAA50519 standard; cDNA; 597 BP.
 XX
 AC AAA50519;
 XX
 DT 05-DEC-2000 (first entry)
 XX
 DE Human cyclin dependent kinase inhibitor p27 DNA.
 XX
 KW Cyclin dependent kinase inhibitor; CDKi; CIP; KIP; human; p27;
 KW angiogenesis; inhibitor; neoplasia; rheumatoid arthritis;
 KW endometriosis; psoriasis; vascular retinopathy; cytostatic;
 KW antiarthritic; antirheumatic; gynaecological; antipsoriatic;
 KW antiproliferative; gene therapy; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200052158-A1.
 XX
 PD 08-SEP-2000.
 XX
 PF 28-FEB-2000; 2000WO-US04970.
 XX
 PR 01-MAR-1999; 99US-0122974.
 PR 05-NOV-1999; 99US-0163682.
 PR 09-DEC-1999; 99US-0457568.
 XX
 PA (CELL-) CELL GENESYS INC.
 PA (MITO-) MITOTIX INC.
 XX
 PI Patel S, McArthur J, Gyuris J;
 XX
 DR WPI; 2000-565501/52.

DR P-PSDB; AAY96066.
 XX Inhibiting angiogenesis and treating angiogenesis-associated
 PT conditions, e.g. neoplasia, psoriasis by transducing an endothelial
 PT cell with a recombinant virus having a transgene encoding a cyclin
 PT dependent kinase inhibitor
 XX
 PS Example 1; Page 127; 138pp; English.
 XX
 CC The present sequence is that of DNA encoding human p27 (see
 CC AAY96066), a cyclin dependent kinase inhibitor (CDK1) that inhibits
 CC angiogenesis. A claimed method for inhibiting angiogenesis
 CC involves transducing an epithelial cell with a transgene encoding
 CC (internalizable, secretable) CDK1. The delivery system for the
 CC transgene may be a liposome or a recombinant virus. The CDK1 is
 CC a protein of the Cip/Kip family such as p27, a protein of the
 CC INK4 family such as p16, active fragments of these proteins (e.g.
 CC amino acids 25-93 or 12-178 of human p27), or a fusion protein
 CC comprising 2 CDK1 proteins such as p27 and p16 (see AAY96068-80).
 CC The method is useful in treating conditions associated with
 CC angiogenesis, e.g. neoplasia, rheumatoid arthritis, endometriosis,
 CC psoriasis and vascular retinopathy (claimed). Alternatively, the
 CC transgene is delivered to an auxiliary cell, and is expressed by
 CC that cell such that the CDK1 is released into the blood and
 CC contacts the target epithelial cell.
 XX
 SQ Sequence 597 BP; 161 A; 164 C; 185 G; 87 T; 0 other;
 Alignment Scores:
 Pred. No.: 0.000172 Length: 597
 Score: 91.00 Matches: 18
 Percent Similarity: 94.74% Conservative: 0
 Best Local Similarity: 94.74% Mismatches: 1
 Query Match: 82.73% Indels: 0
 DB: 21 Gaps: 0
 US-10-042-417A-91 (1-20) x AAY50519 (1-597)
 QY 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
 ID AAY51357
 DB 538 AATGCCGGTCTGTGGAGCAGACGCCCAAGAAGCCTGCCTCAGACAGCTCAACG 594
 RESULT 8
 AAY51357
 ID AAY51357 standard; DNA; 597 BP.
 XX
 AC AAY51357;
 XX
 DT 06-JUN-2000 (first entry)
 XX
 DE Human mutant cyclin-dependent kinase inhibitor (CKI), p27 gene.
 XX
 KW Cyclin-dependent kinase inhibitor; CKI; p27; human; mutant; G1 phase;
 KW KIS; serine/threonine kinase; cell proliferation; modulator; treatment;
 KW cell proliferative disease; vascular disorder; gene therapy; restenosis;
 KW atherosclerosis; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..597
 FT /*tag= a
 FT /product= "Human mutant cyclin-dependent kinase
 FT inhibitor protein, p27"
 FT 28..30
 FT /*tag= b
 FT /note= "ACC to GCT substitution results in a Ser to Ala
 FT change in the amino acid sequence"
 XX
 PN WO200011165-A1.
 XX
 PD 02-MAR-2000.
 XX

PF 20-AUG-1999; . 99WO-US18903.
 XX
 PR 21-AUG-1998; 98US-0057710.
 XX
 PA (NABE/) NABEL G J.
 PA (NABE/) NABEL E 3.
 XX
 PI Nabel GJ, Nabel EG;
 XX
 DR WPI: 2000-237648/20.
 DR P-PSDB; AAY70307.
 XX
 PT Novel serine/threonine kinase hKIS polynucleotides and polypeptides
 PT used for inhibiting the cyclin kinase inhibitor p27, and so alter cell
 PT proliferation
 XX
 PS Example 1; Page 51; 70pp; English.
 XX
 CC The present DNA sequence encodes a mutant cyclin-dependent kinase
 CC inhibitor (CKI), p27. It is bound by hKIS, a serine/threonine kinase,
 CC that inhibits its ability to arrest cells in G1 phase. A substitution
 CC mutation (AGC to GCT) results in a serine to alanine (S10A) change in the
 CC protein, that abolishes phosphorylation of GSR-p27, without affecting
 CC in vitro binding with hKIS. Endogenous p27 was detected in the cytoplasm
 CC and at higher levels in the nucleus. hKIS functions as an inhibitory
 CC kinase of CKI p27. The hKIS sequences are used to modulate cell
 CC proliferation and treat cell proliferative and vascular diseases.
 CC The polynucleotide sequence may be used in gene therapy to treat
 CC vascular disorders such as restenosis or atherosclerosis.
 XX
 SQ Sequence 597 BP; 160 A; 164 C; 185 G; 88 T; 0 other;
 Alignment Scores:
 Pred. No.: 0.000172 Length: 597
 Score: 91.00 Matches: 18
 Percent Similarity: 94.74% Conservative: 0
 Best Local Similarity: 94.74% Mismatches: 1
 Query Match: 82.73% Indels: 0
 DB: 21 Gaps: 0
 US-10-042-417A-91 (1-20) x AAZ51357 (1-597)
 QY 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
 ID AAZ29564
 DB 538 AATGCCGGTCTGTGGAGCAGACGCCCAAGAAGCCTGCCTCAGACAGCTCAACG 594
 RESULT 9
 AAZ29564
 ID AAZ29564 standard; DNA; 597 BP.
 XX
 AC AAZ29564;
 XX
 DT 22-MAR-2000 (first entry)
 XX
 DE Human p27(Kip1) kinase inhibitor protein DNA.
 XX
 KW p27(Kip1) kinase inhibitor protein; FKBP-12; p27(Kip1).FKBP-12 complex;
 KW cytosolic drug-binding protein; yeast two hybrid assay system;
 KW cell differentiation; apoptosis; neurodegeneration; tumorigenicity;
 KW cell proliferation related disorder; atherosclerosis; autoimmune disease;
 KW transplant rejection; inflammation; allergy; cancer; viral infection;
 KW membranous nephropathy; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..597
 FT /*tag= a
 FT /product= "Human p27(Kip1) inhibitor protein"
 FT 137..597
 FT /*tag= b
 FT /note= "Encodes prey sequence that interacts with
 FT FKBP-12"
 XX

```

XX WO9965939-A1.
PN
XX
XX PD
XX PF
XX 23-DEC-1999.
XX
XX 18-JUN-1999; 99WO-US13659.
XX
XX 18-JUN-1998; 98US-0099857.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Nandabalan K, Yang M;
XX
XX WPI; 2000-116763/10.
XX
XX P-PSDB; AAY44400.
XX
XX New complex of p27(kip1) and FKBP-1, for treatment, prevention and
XX diagnosis of, e.g. cancer and autoimmune disease -
XX
XX Claim 10; Fig 1; 78pp; English.
XX
XX The present sequence encodes p27(Kip1) kinase inhibitor protein. This
XX interacts with FKBP-12 (a cytosolic drug-binding protein) to form a
XX p27(kip1).FKBP-12 complex using a modified yeast two hybrid assay system.
XX The nucleic acid sequence encoding the proteins of the complex are used
XX to modulate the functions of the complex and exert a therapeutic
XX effect. This is used in regulating many cell functions, e.g. cell cycle
XX progression, differentiation, apoptosis, neurodegeneration, response to
XX viral infection, tumorigenicity. p27(Kip1).FKBP-12 complex and its
XX corresponding nucleic acid sequence is used in diagnosis and treatment
XX of cell proliferation related disorders. Specified diseases are
XX atherosclerosis; autoimmune diseases (e.g. transplant rejection,
XX inflammation or allergy); neurodegeneration; cancer; membranous
XX nephropathy and viral infections.
XX
XX Sequence 597 BP; 161 A; 164 C; 185 G; 87 T; 0 other;

Alignment Scores:
Pred. No.: 0.000172 Length: 597
Score: 91.00 Matches: 18
Percent Similarity: 94.74% Conservative: 0
Best Local Similarity: 94.74% Mismatches: 1
Query Match: 82.73% Indels: 0
DB: 21 Gaps: 0

US-10-042-417A-91 (1-20) x AA229564 (1-597)

QY 2 AsnAlaGlySerValcIuTrpThrProLysLysProGlyLeuArgArgGlnThr 20
Db 538 AATGCCGGTTCGTGGAGCAGACGCCCAAGAAGACCTGGCTCAGAAGACGTCAAACG 594

RESULT 10
AAH28328
ID AAH28328 standard; DNA; 597 BP.
XX
XX AC AAH28328;
XX
XX 05-SEP-2001 (first entry)
XX
XX Nucleotide sequence of a human p27 KIP1 polypeptide.
XX
XX Knockout mouse; p19 INK4d; p27 KIP1; bradykinase; cell growth;
XX proprioceptive abnormality; neuronal growth; motor disorder;
XX neuronal cell; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 1..597
XX FT /*tag= a
XX FT /product= "p27 KIP1"
XX
XX US6245965-B1.

```

```

XX 12-JUN-2001.
XX
XX 29-JAN-1999; 99US-0240906.
XX
XX 29-JAN-1999; 99US-0240906.
XX
XX (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX
XX Roussel MF, Smeys R, Zindy F, Cunningham J;
XX
XX WPI; 2001-424332/45.
XX
XX P-PSDB; AAB84649.
XX
XX New knockout mouse having a genome comprising a homozygous disruption
XX of both p19 INK4d and p27 KIP1 genes, useful as animal models studying
XX motor disorders having symptoms that include bradykinesia and
XX proprioceptive abnormalities -
XX
XX Disclosure; Columns 31-32; 24pp; English.
XX
XX The present sequence encodes a human p27 KIP1 polypeptide. The
XX specification describes a knockout mouse whose genome is manipulated
XX to comprise a homozygous disruption of both the p19 INK4d and p27 KIP1
XX genes, where homozygous disruption of these genes results in the knockout
XX mouse exhibiting bradykinesia and/or proprioceptive abnormalities, or
XX prevents in the expression of functional p19 INK4d and p27 KIP1 proteins.
XX The knockout mouse and cells may be used to identify potential
XX modulator of cell growth and more particularly neuronal growth. The
XX knockout mouse is useful as animal model for studying motor disorders
XX having symptoms that include bradykinesia and/or proprioceptive
XX abnormalities and/or seizures, and in identifying potential modulators
XX of motor functions. Cells from the knockout mouse may be used as a
XX potential source of differentiated neuronal cells, and for identifying
XX agonists and antagonists of neuronal cell growth.
XX
XX Sequence 597 BP; 161 A; 164 C; 185 G; 87 T; 0 other;

Alignment Scores:
Pred. No.: 0.000172 Length: 597
Score: 91.00 Matches: 18
Percent Similarity: 94.74% Conservative: 0
Best Local Similarity: 94.74% Mismatches: 1
Query Match: 82.73% Indels: 0
DB: 22 Gaps: 0

US-10-042-417A-91 (1-20) x AAH28328 (1-597)

QY 2 AsnAlaGlySerValcIuTrpThrProLysLysProGlyLeuArgArgGlnThr 20
Db 538 AATGCCGGTTCGTGGAGCAGACGCCCAAGAAGACCTGGCTCAGAAGACGTCAAACG 594

RESULT 11
AAC84621
ID AAC84621 standard; DNA; 597 BP.
XX
XX AC AAC84621;
XX
XX 02-APR-2001 (first entry)
XX
XX Human p27 protein encoding DNA.
XX
XX S-phase kinase associated protein; SKP1; SKP2; SKP2-like protein; 2P;
XX CUL-1; cullin; CDC53; p27; cyclin E; Max; Mad; C-Myc; p53; Bax;
XX Bad; Bcl-2; tumour; cytostatic; ds.
XX
XX Homo sapiens.
XX
XX WO200075184-A1.
XX
XX 14-DEC-2000.
XX
XX 05-JUN-2000; 2000WO-US15449.

```

XX PR 04-JUN-1999; 99US-0137494.
 XX PA (UYVA) UNIV YALE.
 XX PI Zhang H, Tsvetkov LM, Kondo T;
 XX XX WPI; 2001-061703/07.
 XX DR P-PSDB; AAB48309.
 XX XX
 XX PT Modulating polypeptide levels in a cell, diagnosing and treating tumor,
 XX PT involves altering levels of proteins such as S-phase kinase associated
 XX PT proteins 1, 2 and cullin/CDC53 proteins -
 XX PS Examples; Page 149; 162pp; English.
 XX XX
 XX CC The invention relates to methods of altering the polypeptide levels in a
 XX CC cell, using proteins selected from S-phase kinase associated proteins 1
 XX CC and 2 (SKP1, SKP2), SKP2-like proteins (ZF) and CUL-1 (a member of the
 XX CC cullin/ CDC53 family of proteins). The method is useful for altering the
 XX CC level of p27, cyclin E, Max, Mad, c-Myc, MDM2, p53, Bax, Bad or Bcl-2
 XX CC polypeptide in a cell. SKP2 and SKP2-like protein levels are useful for
 XX CC detecting tumours, and in monitoring tumor treatment in a mammal. Agents
 XX CC that modulate interactions between SKP and target proteins are useful for
 XX CC treating tumours.
 XX SQ Sequence 597 BP; 161 A; 164 C; 185 G; 87 T; 0 other;
 Alignment Scores:
 Pred. No.: 0.000172 Length: 597
 Score: 91.00 Matches: 18
 Percent Similarity: 94.74% Conservative: 0
 Best Local Similarity: 94.74% Mismatches: 1
 Query Match: 82.73% Indels: 0
 DB: 22 Gaps: 0
 US-10-042-417A-91 (1-20) x AAC84621 (1-597)
 Oy 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
 Db 538 AATCCCGTCTCTGTGGAGCAGACGCCCAAGAGGCTGCCTCAGAGACGTCACACG 594
 RESULT 12
 ABV78025
 ID ABV78025 standard; DNA; 597 BP.
 XX AC ABV78025;
 XX DT 12-NOV-2002 (first entry)
 XX DE Hypoxia-regulated protein coding sequence #45.
 KW Cytostatic; vasotropic; tranquiliser; antiatherosclerotic; gene therapy;
 KW antiinflammatory; vulnerary; gynecological; ophthalmological; vaccine;
 KW hypoxia; tumorigenesis; angiogenesis; apoptosis; cancer;
 KW ischaemic condition; reperfusion injury; retinopathy; neonatal stress;
 KW preclapmsia; atherosclerosis; inflammatory condition; wound healing;
 KW inflammation; erythropoiesis; hair loss; human; gene; ds.
 XX OS Homo sapiens.
 XX XX WO200246465-A2.
 XX PN 13-JUN-2002.
 XX PD 10-DEC-2001; 2001WO-GB05458.
 XX PR 08-DEC-2000; 2000GB-0030076.
 XX PR 08-FEB-2001; 2001GB-0003156.
 XX PR 25-OCT-2001; 2001GB-0025666.
 XX XX (OXFO-) OXFORD BIOMEDICA UK LTD.
 XX PA

PI White J, Mundy CR, Ward NR, Krige D, Kingsman SM, Harris RA;
 PI Rayner WN;
 XX DR WPI; 2002-627238/57.
 XX XX
 XX PT Identifying a gene involved in disease for treating hypoxia-regulated
 XX PT conditions, comprises comparing the transcriptome/proteome of two cell
 XX PT types under different conditions and identifying a differentially
 XX PT regulated gene
 XX XX
 XX PS Claim 37; Page 363; 538pp; English.
 XX CC
 XX CC The present invention relates to methods for identifying genes and
 XX CC proteins that are implicated in a specific disease or physiological
 XX CC condition. The method comprises comparing the transcriptome/proteome of a
 XX CC specialised cell type implicated in a disease or condition with that of a
 XX CC second specialised cell type, under two experimental conditions, and
 XX CC identifying a gene that is differentially regulated in the two
 XX CC specialised cell types under experimental conditions. ABV7873-ABV78116
 XX CC and ABP65061-ABP65257 were identified using the methods of the invention.
 XX CC The coding sequences and proteins are useful for treating a disease in a
 XX CC patient, for manufacture of a medicament for treating hypoxia-regulated
 XX CC conditions, and for regulating tumorigenesis, angiogenesis, apoptosis,
 XX CC biological response to hypoxia conditions, or hypoxic-associated
 XX CC pathology in a patient. The coding sequences and proteins are also useful
 XX CC for monitoring the therapeutic treatment of a disease or physiological
 XX CC condition, such as cancer, ischaemic conditions, reperfusion injury,
 XX CC retinopathy, neonatal stress, preclapmsia, atherosclerosis, inflammatory
 XX CC conditions, wound healing, inflammation, erythropoiesis or hair loss.
 XX SQ Sequence 597 BP; 161 A; 164 C; 185 G; 87 T; 0 other;
 Alignment Scores:
 Pred. No.: 0.000172 Length: 597
 Score: 91.00 Matches: 18
 Percent Similarity: 94.74% Conservative: 0
 Best Local Similarity: 94.74% Mismatches: 1
 Query Match: 82.73% Indels: 0
 DB: 24 Gaps: 0
 US-10-042-417A-91 (1-20) x ABV78025 (1-597)
 Oy 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
 Db 538 AATCCCGTCTCTGTGGAGCAGACGCCCAAGAGGCTGCCTCAGAGACGTCACACG 594
 RESULT 13
 AA172396
 ID AA172396 standard; cDNA; 597 BP.
 XX AC AA172396;
 XX DT 02-MAY-2002 (first entry)
 XX DE p27-Kip1 cDNA.
 XX KW Cell cycle inhibitor; antisense; inner ear; sensory hair cell;
 XX KW support cell; auditory function; hearing disorder;
 XX KW sensory neuronal hearing loss; SNHL; ss.
 XX OS Homo sapiens.
 XX XX
 XX FH Key Location/Qualifiers
 XX FT CDS 1..597
 XX FT /*tag= a
 XX FT /product= "p27-Kip1"
 XX PN WO200204605-A2.
 XX XX 17-JAN-2002.
 XX PD 10-JUL-2001; 2001WO-US21793.
 XX PF
 XX XX

PR 11-JUL-2000; 2000US-0614099.
 XX (OTOG-) OTOGENE USA INC.
 PA (OTOG-) OTOGENE AG.
 XX
 PI Kil J, Gu R, Grigour C, Lowenheim H;
 XX WPI; 2002-171713/22.
 DR P-PSDB; AAB47880.
 XX
 XX Stimulating the formation of inner ear sensory hair cells, useful for
 PT treating hearing disorder involves damaging first inner ear sensory
 PT hair cells and promoting the formation of new sensory hair cells from
 PT inner ear support cells -
 XX
 PS Claim 18; Page 63-64; 77pp; English.
 XX
 CC The sequences given in AAT72395-401 encode cell cycle inhibitors.
 CC These nucleic acids may be hybridised by antisense molecules in the
 CC method of the invention. The method is for stimulating the formation
 CC of an inner ear sensory hair cell from an inner ear support cell and
 CC involves damaging a first inner ear sensory hair cell under conditions
 CC that promote the formation of at least one inner ear sensory hair cell
 CC that is in contact with the damaged first inner ear sensory hair cell.
 CC The method is useful for stimulating the formation of inner ear
 CC cells e.g. sensory hair cells and support cells, for improving an
 CC auditory function in an inner ear, in the treatment of hearing disorder
 CC e.g. sensory neuronal hearing loss (SNHL), to identify genes and/or
 CC proteins that are capable of stimulating the formation of inner ear
 CC sensory hair cells and/or the formation of inner ear support cells
 CC from sensory hair cells. The method damages and/or kills the inner
 CC ear sensory cells, such as sensory hair cells and support cells, which
 CC results in the increased stimulation in the formation of new, inner ear
 CC hair cells, thus resulting in the improved curing of the auditory
 CC function.
 XX
 SQ Sequence 597 BP; 161 A; 164 C; 185 G; 87 T; 0 other;

Alignment Scores:
 Pred. No.: 0.000172 Length: 597
 Score: 91.00 Matches: 18
 Percent Similarity: 94.74% Conservative: 0
 Best Local Similarity: 94.74% Mismatches: 1
 Query Match: 82.73% Indels: 0
 DB: 24 Gaps: 0

US-10-042-417A-91 (1-20) x AAT72396 (1-597)

QY 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
 |||||
 Db 538 AATGCGGTCTGTGGAGCAGACGCCCAAGAAGCTGGCTCAGAGACGTCAAACG 594
 |||||

RESULT 14
 AAT74053
 ID AAT74053 standard; cDNA; 1098 BP.
 XX
 AC AAT74053;
 XX
 DT 16-MAR-1998 (first entry)
 XX
 XX CDK inhibitory fusion protein p16p27 coding sequence.
 DE
 XX
 XX Fusion gene; CDK inhibitor; cyclin-dependant kinase; p27 gene; p16 gene;
 KW chimeric polypeptide; human; binding motif; proliferation control;
 KW cell differentiation; cell-cycle inhibitor; proliferative disorder;
 KW tissue degeneration; therapy; ds.
 XX
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH 1..1098
 FT /*tag= a
 FT
 XX

PN W09727297-A1.
 XX
 PD 31-JUL-1997.
 XX
 PF 17-JAN-1997; 97WO-US00569.
 XX
 PR 23-JAN-1996; 96US-0589981.
 XX
 PA (MITO-) MITOTIX INC.
 XX
 PI Beach D, Gyuris J, Lamphere L;
 XX WPI; 1997-393685/36.
 DR P-PSDB; AAW23536.
 XX
 XX Chimeric inhibitor of cyclin dependent kinase - useful for gene
 PT therapy of cancer and other proliferative and differentiative
 PT diseases
 PS
 XX Claim 13; Page 46-48; 58pp; English.
 XX
 CC This sequence encodes a chimeric polypeptide of the invention. It was
 CC derived from a fusion of the human p27 and p16 cDNA sequences. The
 CC chimeric polypeptides of the invention have cyclin-dependent kinase (CDK)
 CC binding motifs from at least two different proteins that bind to CDKs.
 CC The protein controls proliferation and/or differentiation of cells,
 CC particularly they inhibit cell-cycle progression. They can be used to
 CC treat a wide range of proliferative disorders, e.g. cancer, leukaemia,
 CC psoriasis, atherosclerosis, restenosis, chronic inflammation etc. They
 CC can also treat diseases associated with de-differentiation or
 CC degeneration of tissue, e.g. Alzheimer's, Parkinson's or Huntington's
 CC diseases, gastric ulcers and autoimmune diseases of the peripheral
 CC nervous system. Other applications include reducing growth of hair and
 CC protecting hair follicle cells against cytotoxic treatments, cosmetically
 CC to treat various forms of folliculitis, and to inhibit spermatogenesis or
 CC oogenesis. The chimeric proteins can also be used in vitro to maintain
 CC cells, especially neurons intended for testing specific activity of
 CC trophic factors, at selected points in the cell cycle. The proteins are
 CC more active inhibitors of the CDK/cyclin complex than binding motifs used
 CC individually (since they may bind to CDK involved in different stages of
 CC the cell cycle).
 XX
 SQ Sequence 1098 BP; 234 A; 326 C; 378 G; 160 T; 0 other;

Alignment Scores:
 Pred. No.: 0.000341 Length: 1098
 Score: 91.00 Matches: 18
 Percent Similarity: 94.74% Conservative: 0
 Best Local Similarity: 94.74% Mismatches: 1
 Query Match: 82.73% Indels: 0
 DB: 18 Gaps: 0

US-10-042-417A-91 (1-20) x AAT74053 (1-1098)

QY 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
 |||||
 Db 1039 AATGCGGTCTGTGGAGCAGACGCCCAAGAAGCTGGCTCAGAGACGTCAAACG 1095
 |||||

RESULT 15
 AAX26224
 ID AAX26224 standard; cDNA; 1098 BP.
 XX
 AC AAX26224;
 XX
 DT 25-MAY-1999 (first entry)
 XX
 XX Human p16p27 fusion protein encoding cDNA.
 DE
 XX
 KW Cyclin-dependent kinase; CDK; CDK/cyclin complex; inhibitory; restenosis;
 KW CDK-binding motif; endothelialisation; fusion protein; therapeutic; acne;
 KW intracellular; transcellular; transcytosis; vascular wound; repair; hair;
 KW smooth muscle; cardiovascular; arteriosclerotic; fibrotic disorder;
 KW cellular proliferation; rheumatoid arthritis; diabetes; cirrhosis; graft;

KW tumour; inflammation; neurodegeneration; periodontal; spermatogenesis;
 XX tachycardia; human; p27; p16; ds.

OS Homo sapiens.

XX
 FH Key Location/Qualifiers
 FT CDS 1..1098
 FT /*tag= a

XX WO9906540-A2.

XX
 PD 11-FEB-1999.

XX 29-JUL-1998; 98WO-US15759.

XX 29-JUL-1997; 97US-0902572.

XX (MITO-) MITOTIF INC.

XX Beach DH, Gyuris J, Lamphere L;

XX WPI; 1999-153770/13.

XX P-PSDB; AAW95096.

XX Fusion and chimaeric proteins including cyclin-dependent kinase
 PT binding motif - used for regulation of cell proliferation and
 PT differentiation, for treatment of, e.g. vascular injury, cancers,
 PT fibrosis and neurodegeneration

XX Claim 60; Page 78-79; 88pp; English.

XX The invention relates to novel inhibitors of cyclin-dependent kinases
 CC (CDKs), particularly CDK/cyclin complexes. It provides a recombinant
 CC transfection system (A) that comprises: (i) first gene construct
 CC comprising a sequence encoding an inhibitory polypeptide containing at
 CC least one CDK-binding motif for binding and inhibiting activity of a
 CC CDK, linked to a transcription regulator functional in eukaryotic cells;
 CC (ii) second gene construct comprising a sequence encoding a polypeptide
 CC that promotes endotheialisation, and (iii) a gene delivery composition
 CC for delivering the GCs to a cell for transfection. Also provided are
 CC nucleic acids encoding a fusion protein (FP) containing: (i) a
 CC therapeutic polypeptide sequence (TP) from an intracellular protein that
 CC alters a cellular process when FP enters the cell, and (ii) a
 CC transcellular polypeptide sequence (TCP) that promotes transcytosis of
 CC FP. The FP consists of at least one CDK-binding motif and a TCP. See
 CC AAX26220 for detailed uses of the recombinant transfection system. The
 CC present sequence represents a cDNA encoding a human p16p27 fusion
 CC protein.

XX SQ Sequence 1098 BP; 234 A; 326 C; 378 G; 160 T; 0 other;

Alignment Scores:
 Pred. No.: 0.000341 Length: 1098
 Score: 91.00 Matches: 18
 Percent Similarity: 94.74% Conservative: 0
 Best Local Similarity: 94.74% Mismatches: 1
 Query Match: 82.73% Indels: 0
 DB: 20 Gaps: 0

US-10-042-417A-91.(1-20) x AAX26224 (1-1098)

QY 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
 DB 1039 AATGCCGGTCTCTGTGAGCAGAGCCCAAGAGCCCTGGCTCAGAGACGTCACACG 1095

Search completed: August 28, 2003, 13:47:45
 Job time : 288 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: August 28, 2003, 13:29:25 ; Search time 2746 Seconds
(without alignments)
297.958 Million cell updates/sec

Title: US-10-042-417A-91
Perfect score: 110
Sequence: 1 NNAGSVETPKPKGLRRQT 20

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 2045481386 residues
Total number of hits satisfying chosen parameters: 5777422
Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPTO.spool/US10042417/runat_19082003_141154_27886/app.query.fasta_1.199
-DB=GenEmbl -GFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10042417.ecgn_1.1.3508.erunat_19082003_141154_27886 -NCPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -YGAPOP=10 -XGAPOP=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_in.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*

29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rod.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	91	82.7	327	9	S76986S2	S76988 p27Kipl-cyc
2	91	82.7	596	6	AR200007	AR200007 Sequence
3	91	82.7	597	6	AR000109	AR000109 Sequence
4	91	82.7	597	6	AR125495	AR125495 Sequence
5	91	82.7	597	6	AR157884	AR157884 Sequence
6	91	82.7	597	6	AR219459	AR219459 Sequence
7	91	82.7	597	6	AR282523	AR282523 Sequence
8	91	82.7	597	6	AX057188	AX057188 Sequence
9	91	82.7	597	6	AX376623	AX376623 Sequence
10	91	82.7	597	9	AF247551	AF247551 Homo sapi
11	91	82.7	597	9	HSU10906	U10906 Human cycli
12	91	82.7	948	4	D84649	D84649 Felis catus
13	91	82.7	1098	6	AR219451	AR219451 Sequence
14	91	82.7	1098	6	AR267147	AR267147 Sequence
15	91	82.7	1098	6	AR282515	AR282515 Sequence
16	91	82.7	1098	6	BD080713	BD080713 Inhibitor
17	91	82.7	1121	6	AR219449	AR219449 Sequence
18	91	82.7	1121	6	AR282513	AR282513 Sequence
19	91	82.7	1121	6	BD080724	BD080724 Inhibitor
20	91	82.7	1143	6	AR219450	AR219450 Sequence
21	91	82.7	1143	6	AR267146	AR267146 Sequence
22	91	82.7	1143	6	AR282514	AR282514 Sequence
23	91	82.7	1143	6	BD080712	BD080712 Inhibitor
24	91	82.7	1420	6	AR219448	AR219448 Sequence
25	91	82.7	1420	6	AR267143	AR267143 Sequence
26	91	82.7	1420	6	AR282512	AR282512 Sequence
27	91	82.7	1420	6	BD080709	BD080709 Inhibitor
28	91	82.7	1420	6	I67718	I67718 Sequence 1
29	91	82.7	2090	9	AY004255	AY004255 Homo sapi
30	91	82.7	2334	9	BC001971	BC001971 Homo sapi
31	91	82.7	5727	9	AF480891	AF480891 Homo sapi
32	91	82.7	157892	9	AC008115	AC008115 Homo sapi
33	86	78.2	714	10	CGU49649	U49649 Cricetus
34	82	74.5	1327	4	AB031957	AB031957 Sus scrof
35	82	74.5	1501	4	AB031956	AB031956 Sus scrof
36	82	74.5	1958	4	AB031955	AB031955 Sus scrof
37	82	74.5	1958	6	BD017076	BD017076 Nucleic a
38	78	70.9	236	9	HS48C11F	Z61220 H.sapiens C
39	74	67.3	199	6	AX524640	AX524640 Sequence
40	74	67.3	593	6	AX553378	AX553378 Sequence
41	74	67.3	593	6	AR200008	AR200008 Sequence
42	74	67.3	594	6	AR157885	AR157885 Sequence
43	74	67.3	594	10	AF015194	AF015194 Rattus no
44	74	67.3	594	10	D86924	D86924 Rattus norv
45	74	67.3	594	10	MMU09968	U09968 Mus musculu

ALIGNMENTS

```
S76986S2
LOCUS       S76986S2               327 bp    DNA        linear    PRI 26-SEP-1995
DEFINITION  p27Kipl-cyclin-dependent kinase inhibitor [human, fetal brain,
Genomic, 327 nt, segment 2 of 2].
ACCESSION   S76986
KEYWORDS    S76988.1  GI:998402
SOURCE      2 of 2
            Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 327)
AUTHORS     Pietenpol,J.A., Bohlander,S.K., Sato,Y., Papadopoulos,N., Liu,B.,
            Friedman,C., Trask,B.J., Roberts,J.M., Kinzler,K.W., Rowley,J.D.
            et al.
TITLE       Assignment of the human p27Kipl gene to 12p13 and its analysis in
            leukemias
JOURNAL     Cancer Res. 55 (6), 1206-1210 (1995)
MEDLINE     95188144
PUBMED      7882309
REMARK      GenBank staff at the National Library of Medicine created this
            entry [NCBI gibbsq 164153] from the original journal article.
            This sequence comes from Table 1.
            Map location: 12p13.
            Authors indicate codon 109 different from previously published
            sequence.
FEATURES             Location/Qualifiers
     source          1..327
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     mRNA            join(S76986.1:1..501:99..>228)
     gene            order(S76986.1:27..596:1..220)
                     /gene="p27Kipl"
                     /note="cyclin-dependent kinase inhibitor"
     CDS             join(S76986.1:27..501:99..220)
                     /gene="p27Kipl"
                     /note="cyclin-dependent kinase inhibitor"
                     /codon_start=1
                     /protein_id="AAD14244.1"
                     /db_xref="GI:4261944"
     translation     /translation="MSNVVSVNGSPSLERMDARQADHPKPSACRNLFQVDHEELTRD
            LEKCHRMEEASQKWNDFQNHKPLEYQWEKESLPEFYVRPPRPGACKVP
            AQESODGSGSRPAAPLIGAPANSEDTLHVDKTPDSOTGLAEQCAGTKRKPATDTS
            STQNKRAANTEENVSDGSPNAGSVETPKPKGLRRQT"
BASE COUNT        90 a 79 c 74 g 84 t
ORIGIN
Alignment Scores:
Pred. No.: 0.000219 Length: 327
Score: 91.00 Matches: 18
Percent Similarity: 94.74% Conservative: 0
Best Local Similarity: 94.74% Mismatches: 1
Query Match: 82.73% Indels: 0
DB: 9 Gaps: 0

US-10-042-417A-91 (1-20) x S76986S2 (1-327)

QY      2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
      |||||||
Db      161 AATGCCGGTCTGTGGAGCAGAGCCCAAGAGCGCTGCGCTCAGAAGACGTCAAACG 217

RESULT 2
LOCUS       AR200007               596 bp    DNA        linear    PAT 20-APR-2002
DEFINITION  Sequence 1 from patent US 6355774.
ACCESSION   AR200007
KEYWORDS    AR200007.1  GI:20250081
SOURCE      Unknown.
            Nabel,G.J. and Nabel,E.G.
            Method for treating vascular proliferative diseases with p27 and
            fusions thereof.
            Unclassified.

US-10-042-417A-91 (1-20) x S76986S2 (1-327)

QY      2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
      |||||||
Db      161 AATGCCGGTCTGTGGAGCAGAGCCCAAGAGCGCTGCGCTCAGAAGACGTCAAACG 217

RESULT 4
LOCUS       AR125495               597 bp    DNA        linear    PAT 16-MAY-2001
DEFINITION  Sequence 1 from patent US 6177272.
ACCESSION   AR125495
KEYWORDS    AR125495.1  GI:14111557
SOURCE      Unknown.
            Unclassified.
            1 (bases 1 to 597)
            Nabel,G.J. and Nabel,E.G.
            Method for treating vascular proliferative diseases with p27 and
            fusions thereof.
            Unclassified.

US-10-042-417A-91 (1-20) x AR000109 (1-597)

QY      2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
      |||||||
Db      538 AATGCCGGTCTGTGGAGCAGAGCCCAAGAGCGCTGCGCTCAGAAGACGTCAAACG 594

RESULT 3
LOCUS       AR000109               597 bp    DNA        linear    PAT 04-DEC-1998
DEFINITION  Sequence 3 from patent US 5736318.
ACCESSION   AR000109
KEYWORDS    AR000109.1  GI:3962640
SOURCE      Unknown.
            Unclassified.
            1 (bases 1 to 597)
            Munger,K. and Jones,D.Leanne.
            Method and kit for evaluating human papillomavirus transformed
            cells
            Patent: US 5736318-A 3 07-APR-1998;
            Location/Qualifiers
            source          1..597
            /organism="unknown"
BASE COUNT        161 a 164 c 185 g 87 t
ORIGIN
Alignment Scores:
Pred. No.: 0.000392 Length: 597
Score: 91.00 Matches: 18
Percent Similarity: 94.74% Conservative: 0
Best Local Similarity: 94.74% Mismatches: 1
Query Match: 82.73% Indels: 0
DB: 6 Gaps: 0

US-10-042-417A-91 (1-20) x AR200007 (1-596)

QY      2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
      |||||||
Db      538 AATGCCGGTCTGTGGAGCAGAGCCCAAGAGCGCTGCGCTCAGAAGACGTCAAACG 594

RESULT 3
LOCUS       AR000109               597 bp    DNA        linear    PAT 04-DEC-1998
DEFINITION  Sequence 3 from patent US 5736318.
ACCESSION   AR000109
KEYWORDS    AR000109.1  GI:3962640
SOURCE      Unknown.
            Unclassified.
            1 (bases 1 to 597)
            Munger,K. and Jones,D.Leanne.
            Method and kit for evaluating human papillomavirus transformed
            cells
            Patent: US 5736318-A 3 07-APR-1998;
            Location/Qualifiers
            source          1..597
            /organism="unknown"
BASE COUNT        161 a 164 c 185 g 87 t
ORIGIN
Alignment Scores:
Pred. No.: 0.000392 Length: 597
Score: 91.00 Matches: 18
Percent Similarity: 94.74% Conservative: 0
Best Local Similarity: 94.74% Mismatches: 1
Query Match: 82.73% Indels: 0
DB: 6 Gaps: 0

US-10-042-417A-91 (1-20) x AR000109 (1-597)

QY      2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
      |||||||
Db      538 AATGCCGGTCTGTGGAGCAGAGCCCAAGAGCGCTGCGCTCAGAAGACGTCAAACG 594

RESULT 4
LOCUS       AR125495               597 bp    DNA        linear    PAT 16-MAY-2001
DEFINITION  Sequence 1 from patent US 6177272.
ACCESSION   AR125495
KEYWORDS    AR125495.1  GI:14111557
SOURCE      Unknown.
            Unclassified.
            1 (bases 1 to 597)
            Nabel,G.J. and Nabel,E.G.
            Method for treating vascular proliferative diseases with p27 and
            fusions thereof.
            Unclassified.
```

JOURNAL Patent: US 6177272-A 1 23-JAN-2001;

FEATURES Location/Qualifiers

source 1..597

BASE COUNT 161 a 164 c 185 g 87 t

ORIGIN

Alignment Scores:
 Pred. No.: 0.000392 Length: 597
 Score: 91.00 Matches: 18
 Percent Similarity: 94.74% Conservative: 0
 Best Local Similarity: 94.74% Mismatches: 1
 Query Match: 82.73% Indels: 0
 DB: 6 Gaps: 0

US-10-042-417A-91 (1-20) x AR125495 (1-597)

QY 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
 DB 538 AATGCCGTTCTGTGGAGCAGACGCCCAAGAAGCCTGGCCTCAGAAGACGTCAAACG 594

RESULT 5

LOCUS AR157884 597 bp DNA linear PAT 17-OCT-2001

DEFINITION Sequence 5 from patent US 6245965.

ACCESSION AR157884

VERSION AR157884.1 GI:16218901

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 597)

AUTHORS Roussel,M.F., Sneyne,R., Zindy,F. and Cunningham,J.

TITLE Knockout mice and cells that lack p19INK4d and p27KIP1 activity and methods of use thereof

JOURNAL Patent: US 6245965-A 5 12-JUN-2001;

FEATURES Location/Qualifiers

source 1..597

BASE COUNT 161 a 164 c 185 g 87 t

ORIGIN

Alignment Scores:
 Pred. No.: 0.000392 Length: 597
 Score: 91.00 Matches: 18
 Percent Similarity: 94.74% Conservative: 0
 Best Local Similarity: 94.74% Mismatches: 1
 Query Match: 82.73% Indels: 0
 DB: 6 Gaps: 0

US-10-042-417A-91 (1-20) x AR157884 (1-597)

QY 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
 DB 538 AATGCCGTTCTGTGGAGCAGACGCCCAAGAAGCCTGGCCTCAGAAGACGTCAAACG 594

RESULT 6

LOCUS AR219459 597 bp DNA linear PAT 25-SEP-2002

DEFINITION Sequence 25 from patent US 6420345.

ACCESSION AR219459

VERSION AR219459.1 GI:23320626

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 597)

AUTHORS Patel,S., McArthur,J. and Gyuris,J.

TITLE Methods and reagents for inhibiting angiogenesis

JOURNAL Patent: US 6420345-A 25 16-JUL-2002;

FEATURES Location/Qualifiers

source 1..597

/organism="unknown"

BASE COUNT 161 a 164 c 185 g 87 t

ORIGIN

Alignment Scores:
 Pred. No.: 0.000392 Length: 597
 Score: 91.00 Matches: 18
 Percent Similarity: 94.74% Conservative: 0
 Best Local Similarity: 94.74% Mismatches: 1
 Query Match: 82.73% Indels: 0
 DB: 6 Gaps: 0

US-10-042-417A-91 (1-20) x AR219459 (1-597)

QY 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
 DB 538 AATGCCGTTCTGTGGAGCAGACGCCCAAGAAGCCTGGCCTCAGAAGACGTCAAACG 594

RESULT 7

LOCUS AR282523

DEFINITION Sequence 15 from patent US 6521602.

ACCESSION AR282523

VERSION AR282523.1 GI:29718997

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 597)

AUTHORS Patel,S., McArthur,J., Gyuris,J., Mendez,M.J. and Finer,M.H.

TITLE Anti-neoplastic compositions and uses thereof

JOURNAL Patent: US 6521602-A 25 18-FEB-2003;

FEATURES Location/Qualifiers

source 1..597

/organism="unknown"

BASE COUNT 161 a 164 c 185 g 87 t

ORIGIN

Alignment Scores:
 Pred. No.: 0.000392 Length: 597
 Score: 91.00 Matches: 18
 Percent Similarity: 94.74% Conservative: 0
 Best Local Similarity: 94.74% Mismatches: 1
 Query Match: 82.73% Indels: 0
 DB: 6 Gaps: 0

US-10-042-417A-91 (1-20) x AR282523 (1-597)

QY 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
 DB 538 AATGCCGTTCTGTGGAGCAGACGCCCAAGAAGCCTGGCCTCAGAAGACGTCAAACG 594

RESULT 8

LOCUS AX057188

DEFINITION Sequence 64 from Patent WO0075184.

ACCESSION AX057188

VERSION AX057188.1 GI:12309990

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1

AUTHORS Zhang,H., Tsvetkov,L.M. and Kondo,T.

TITLE Modulation of protein levels using the scf complex

JOURNAL Patent: WC 0075184-A 64 14-DEC-2000;

FEATURES Location/Qualifiers

source 1..597

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

BASE COUNT 161 a 164 c 185 g 87 t

ORIGIN

AX057188 Sequence 64 from Patent WO0075184.
 AX057188
 AX057188.1 GI:12309990
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 Zhang,H., Tsvetkov,L.M. and Kondo,T.
 Modulation of protein levels using the scf complex
 Patent: WC 0075184-A 64 14-DEC-2000;
 YALE UNIVERSITY (US)
 Location/Qualifiers
 1..597
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

```

ORIGIN
Alignment Scores:
Pred. No.: 0.000392 Length: 597
Score: 91.00 Matches: 18
Percent Similarity: 94.74% Conservative: 0
Best Local Similarity: 94.74% Mismatches: 1
Query Match: 82.73% Indels: 0
DB: 6 Gaps: 0

US-10-042-417A-91 (1-20) x AX057188 (1-597)
Qy 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
|||||
Db 538 AATGCCGGTCTGTGGAGCAGACGCCCAAGAAGCTGGCCTCAGAACGCTCAAACG 594

RESULT 9
AX376623
LOCUS AX376623 597 bp DNA linear PAT 01-MAR-2002
DEFINITION Sequence 8 from Patent WO0204605.
ACCESSION AX376623
VERSION AX376623.1 GI:19170710
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Kil, J., Gu, R., Griguer, C. and Lowenheim, H.
TITLE Stimulation of cellular regeneration and differentiation in the
inner ear
JOURNAL Patent: WO 0204605-A 8 17-JAN-2002;
Otogene USA, Inc. (US); Otogene, AG (DE)
FEATURES
source
1. 597
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
1. 597
/feature="unnamed protein product"
/codon_start=1
/protein_id="CAD26826.1"
/db_xref="GI:19170711"
/translaton="MSNVRVSNGLSPSLERMDARQAEHPKPSACRNLFPGVDHEELTRD
LEKRCRDEEASQRKNWFQNHKPLEGKYEQVKSLEPFEYRPPRPKGAACKVP
AQESODVSGSRPAAPLIGAPANSEDTHLVDPKTPDSQTLAEQACAGIRKRPATDSS
STONKRNARTEENVSDGSPNAGSVEQTPPKPLRRRT"
BASE COUNT 161 a 164 c 185 g 87 t
ORIGIN

Alignment Scores:
Pred. No.: 0.000392 Length: 597
Score: 91.00 Matches: 18
Percent Similarity: 94.74% Conservative: 0
Best Local Similarity: 94.74% Mismatches: 1
Query Match: 82.73% Indels: 0
DB: 6 Gaps: 0

US-10-042-417A-91 (1-20) x AF247551 (1-597)
Qy 2 AsnAlaGlySerValGluTrpThrProLysLysProGlyLeuArgArgGlnThr 20
|||||
Db 538 AATGCCGGTCTGTGGAGCAGACGCCCAAGAAGCTGGCCTCAGAACGCTCAAACG 594

RESULT 11
HSU10906
LOCUS HSU10906 597 bp mRNA linear PRI 27-JUL-1994
DEFINITION Human cyclin-dependent kinase inhibitor p27kip1 mRNA, complete cds.
ACCESSION U10906
VERSION U10906.1 GI:516558
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Polyak, K., Lee, M., Hong, E., Erdjument-Bromage, H., Koff, A.,
Roberts, J. M., Tempst, P. and Massague, J.
TITLE Cloning of p27kip1, a cyclin-dependent kinase inhibitor, and a
potential mediator of extracellular antimitogenic signals
JOURNAL Cell 78, 56-66 (1994)
REFERENCE 2
AUTHORS Massague, J.
TITLE Direct Submission
JOURNAL Submitted (15-JUN-1994) M.-H. Lee and J. Massague, Cell Biology and
Genetics, Memorial Sloan-Kettering Cancer Center, 1275 York Avenue,
New York, NY 10021, USA
FEATURES
source
1. 597
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

```

CDS
/clone="p27 kipl"
/sex="male"
/tissue_type="kidney"
/dev_stage="adult"
1..597
/function="cyclin-dependent kinase inhibitor"
/codon_start=1
/product="p27kipl"
/protein_id="AAA20240.1"
/db_xref="GI:516559"
/translation="MSNVRVSGSPSLERMDARQAEHPKPSACRNLFPGVPDHEELTRD
LEKHDRMERASORKWNFDQNHKPLEGKYEWQVEKGSLEFYYRPPRPKACKVP
AQESQDVSGNRPAAPLIPANSEDTHLVDKTPDSQTLAEQCCAGIRKRPATDGS
STQNKRRNTEENVSDGSPNAGSVQETPKPKPLRRROT"
BASE COUNT 161 a 164 c 185 g 87 t
ORIGIN
Alignment Scores:
Pred. No.: 0.000392 Length: 597
Score: 91.00 Matches: 18
Percent Similarity: 94.74% Conservative: 0
Best Local Similarity: 94.74% Mismatches: 1
Query Match: 82.73% Indels: 0
DB: 9 Gaps: 0
US-10-042-417A-91 (1-20) x HSU10906 (1-597)
Qy 2 AsnAlaGlySerValGluTptThrProLysLysProGlyLeuArgArgGlnThr 20
|||||
Db 538 AATGCCGGTCTGTGGAGCAGACGCCCAAGAGCGCTGCCTCAGACAGCTCAACG 594
RESULT 12
D84649 948 bp mRNA linear MAM 31-OCT-1997
LOCUS
DEFINITION Felis catus mRNA for p27/Kipl, complete cds.
ACCESSION D84649
VERSION D84649.1 GI:258786
KEYWORDS p27/Kipl.
SOURCE Felis catus (cat)
ORGANISM Felis catus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
REFERENCE 1 (sites)
AUTHORS Okuda,M., Minehata,K., Setoguchi,A., Cho,K.-W., Nakamura,N.,
Nishigaki,K., Wataai,T., Cevalario,S., O'Brien,S.J., Tsujimoto,H. and
Hasegawa,A.
TITLE Cloning and chromosome mapping of the feline genes p21WAF1 and
p27Kipl
JOURNAL Gene 198 (1-2), 141-147 (1997)
MEDLINE 98036042
PUBMED 9370275
REFERENCE 2 (bases 1 to 948)
AUTHORS Okuda,M.
TITLE Direct Submission
JOURNAL Submitted (26-APR-1996) Masaru Okuda, The University of Tokyo,
Veterinary Internal Medicine; Yayoi 1-1-1, Bunkyo-ku, Tokyo 113,
Japan (E-mail:aa4711@hongo.ecc.u-tokyo.ac.jp,
Tel:03-3812-2111(ex.5403), Fax:03-5800-6866)
FEATURES
source
1..948
/organism="Felis catus"
/mol_type="mRNA"
/db_xref="taxon:9685"
/clone="pHFK.5"
/tissue_type="Lymph node"
62..658
/codon_start=1
/product="p27/Kipl"
/protein_id="BAA23167.1"
/db_xref="GI:2588787"
/translation="MSNVRVSGSPSLERMDARQAEHPKPSACRNLFPGVPDHEELTRD
LEKHDRMERASORKWNFDQNHKPLEGKYEWQVEKGSLEFYYRPPRPKACKVP
AQESQDVSGNRPAAPLIPANSEDTHLVDKTPDSQTLAEQCCAGIRKRPATDGS

BASE COUNT 277 a 222 c 248 g 201 t
ORIGIN
Alignment Scores:
Pred. No.: 0.000612 Length: 948
Score: 91.00 Matches: 18
Percent Similarity: 94.74% Conservative: 0
Best Local Similarity: 94.74% Mismatches: 1
Query Match: 82.73% Indels: 0
DB: 4 Gaps: 0
US-10-042-417A-91 (1-20) x D84649 (1-948)
Qy 2 AsnAlaGlySerValGluTptThrProLysLysProGlyLeuArgArgGlnThr 20
|||||
Db 599 AACGCTGGTTCAGTGGAGCAGACGCCCAAGAGCGCTGCCTCAGACAGCTCAACG 655
RESULT 13
AR219451
LOCUS
DEFINITION Sequence 9 from patent US 6420345.
ACCESSION AR219451
VERSION AR219451.1 GI:23320618
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1098)
AUTHORS Patel,S., McArthur,J. and Gyuris,J.
TITLE Methods and reagents for inhibiting angiogenesis
JOURNAL Patent: US 6420345-A 9 16-JUL-2002;
FEATURES Location/Qualifiers
1..1098
source /organism="unknown"
BASE COUNT 234 a 326 c 378 g 160 t
ORIGIN
Alignment Scores:
Pred. No.: 0.000706 Length: 1098
Score: 91.00 Matches: 18
Percent Similarity: 94.74% Conservative: 0
Best Local Similarity: 94.74% Mismatches: 1
Query Match: 82.73% Indels: 0
DB: 6 Gaps: 0
US-10-042-417A-91 (1-20) x AR219451 (1-1098)
Qy 2 AsnAlaGlySerValGluTptThrProLysLysProGlyLeuArgArgGlnThr 20
|||||
Db 1039 AATGCCGGTCTGTGGAGCAGACGCCCAAGAGCGCTGCCTCAGACAGCTCAACG 1095
RESULT 14
AR267147
LOCUS
DEFINITION Sequence 7 from patent US 6495526.
ACCESSION AR267147
VERSION AR267147.1 GI:29696937
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1098)
AUTHORS Gyuris,J., Lamphere,L. and Beach,D.H.
TITLE Inhibitors of cell-cycle progression and uses related thereto
JOURNAL Patent: US 6495526-A 7 17-DEC-2002;
FEATURES Location/Qualifiers
1..1098
source /organism="unknown"
BASE COUNT 234 a 326 c 378 g 160 t
ORIGIN
Alignment Scores:

Pred. No.: 0.000706 Length: 1098
 Score: 91.00 Matches: 18
 Percent Similarity: 94.74% Conservative: 0
 Best Local Similarity: 94.74% Mismatches: 1
 Query Match: 82.73% Indels: 0
 DB: 6 Gaps: 0

US-10-042-417A-91 (1-20) x AR267147 (1-1098)

Qy 2 AsnAlaGlySerValGluThrProLysLysProGlyLeuArgArgGlnThr 20
 Db 1039 AATGCCGGTTCTGTGGAGCAGACGCCCAAGAAAGCTGGCCTCAGAGACGCTCAAACG 1095

RESULT 15

AR282515
 LOCUS AR282515 1098 bp DNA linear PAT 10-APR-2003
 DEFINITION Sequence 9 from patent US 6521602.
 ACCESSION AR282515
 VERSION AR282515.1 GI:29718989
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 1098)
 AUTHORS Patel,S., McArthur,J., Gyuris,J., Mendez,M.J. and Finer,M.H.
 TITLE Anti-neoplastic compositions and uses thereof
 JOURNAL Patent: US 6521602-A 9 18-FEB-2003;
 FEATURES Location/Qualifiers
 source 1..1098

BASE COUNT 234 a 326 c 378 g 160 t
 ORIGIN

Alignment Scores:

Pred. No.: 0.000706 Length: 1098
 Score: 91.00 Matches: 18
 Percent Similarity: 94.74% Conservative: 0
 Best Local Similarity: 94.74% Mismatches: 1
 Query Match: 82.73% Indels: 0
 DB: 6 Gaps: 0

US-10-042-417A-91 (1-20) x AR282515 (1-1098)

Qy 2 AsnAlaGlySerValGluThrProLysLysProGlyLeuArgArgGlnThr 20
 Db 1039 AATGCCGGTTCTGTGGAGCAGACGCCCAAGAAAGCTGGCCTCAGAGACGCTCAAACG 1095

Search completed: August 28, 2003, 14:33:48
 Job time : 2751 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2003, 09:42:03 ; Search time 21 Seconds
(without alignments)
27.508 Million cell updates/sec

Title: US-10-042-417A-91

Perfect score: 110
Sequence: 1 NNAGSVEWTPKKPGLRRROT 20Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 138335 seqs, 28883092 residues

Total number of hits satisfying chosen parameters: 138335

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_New.*
1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep.*
2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep.*
3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep.*
4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep.*
5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep.*
6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep.*
7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	110	100.0	20	6	US-10-042-417A-91
2	91	82.7	19	6	US-10-042-417A-85
3	91	82.7	198	7	US-60-490-890-717
4	47	42.7	443	6	US-10-408-765A-2578
5	44.5	40.5	1101	6	US-10-408-765A-2118
6	44.5	40.5	1110	7	US-60-490-890-2574
7	43	39.1	238	6	US-10-383-834-4
8	43	39.1	367	7	US-60-494-568-17
9	43	39.1	2861	6	US-10-374-979-108
10	43	39.1	2861	6	US-10-331-496A-89
11	42	38.2	57	6	US-10-273-573-10183
12	41.5	37.7	271	1	PCT-US03-23249-34
13	41.5	37.7	435	6	US-10-467-042-8
14	41	37.3	95	6	US-10-308-817-103
15	41	37.3	96	6	US-10-308-817-107
16	41	37.3	196	6	US-10-273-573-7911
17	41	37.3	307	6	US-10-612-783-3897
18	40.5	36.8	324	5	US-09-897-516A-8023
19	40	36.4	300	6	US-10-603-113-18445
20	40	36.4	750	6	US-10-612-783-5980
21	40	36.4	766	6	US-10-612-783-5979
22	39.5	35.9	524	6	US-10-603-113-18975
23	39	35.5	65	6	US-10-603-113-25320
24	39	35.5	308	5	US-09-897-516A-7040
25	39	35.5	320	6	US-10-603-113-15469
26	39	35.5	377	6	US-10-293-244-3540

27	39	35.5	394	5	US-09-897-516A-7628	Sequence 7628, Ap
28	39	35.5	394	6	US-10-273-573-7259	Sequence 7259, Ap
29	39	35.5	406	6	US-10-293-244-1572	Sequence 1572, Ap
30	39	35.5	483	6	US-10-286-897-5699	Sequence 5699, Ap
31	39	35.5	483	6	US-10-258-898A-5699	Sequence 5699, Ap
32	39	35.5	490	5	US-09-674-546A-2641	Sequence 2641, Ap
33	39	35.5	490	5	US-09-674-546A-2643	Sequence 2643, Ap
34	39	35.5	569	6	US-10-273-573-8959	Sequence 8959, Ap
35	39	35.5	1037	6	US-10-603-114-7677	Sequence 7677, Ap
36	39	35.5	1164	5	US-09-897-516A-8040	Sequence 8040, Ap
37	39	35.5	2234	1	PCT-US03-20907-20	Sequence 20, Appl
38	39	35.5	2234	6	US-10-612-090-20	Sequence 20, Appl
39	38.5	35.0	275	7	US-60-487-610-1729	Sequence 1729, Ap
40	38.5	35.0	275	7	US-60-487-610-1855	Sequence 1855, Ap
41	38.5	35.0	282	7	US-60-487-610-1728	Sequence 1728, Ap
42	38.5	35.0	333	1	PCT-US03-23249-61	Sequence 61, Appl
43	38.5	35.0	696	1	PCT-US02-29560A-235	Sequence 235, App
44	38	34.5	69	6	US-10-631-402-2084	Sequence 2084, Ap
45	38	34.5	69	6	US-10-631-441-2084	Sequence 2084, Ap

ALIGNMENTS

RESULT 1
US-10-042-417A-91
; Sequence 91, Application US/10042417A
; GENERAL INFORMATION:
; APPLICANT: Pagano, M.
; TITLE OF INVENTION: METHODS TO IDENTIFY COMPOUNDS USEFUL FOR THE TREATMENT OF PROL
; FILE REFERENCE: 591A-090-999
; CURRENT APPLICATION NUMBER: US/10/042,417A
; CURRENT FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: 60/260,179
; PRIOR FILING DATE: 2001-01-05
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 91
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: Phosphorylation
; LOCATION: 9
; OTHER INFORMATION: Phosphothreonine
US-10-042-417A-91

Query Match 100.0%; Score 110; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.5e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NNAGSVEWTPKKPGLRRROT 20
Db 1 NNAGSVEWTPKKPGLRRROT 20
|||||

RESULT 2
US-10-042-417A-85
; Sequence 85, Application US/10042417A
; GENERAL INFORMATION:
; APPLICANT: Pagano, M.
; TITLE OF INVENTION: METHODS TO IDENTIFY COMPOUNDS USEFUL FOR THE TREATMENT OF PROL
; FILE REFERENCE: 591A-090-999
; CURRENT APPLICATION NUMBER: US/10/042,417A
; CURRENT FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: 60/260,179
; PRIOR FILING DATE: 2001-01-05
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 85
; LENGTH: 19

RESULT 7
US-10-383-834-4
; Sequence 4: Application US/10383834

; GENERAL INFORMATION:
; APPLICANT: Chang, Chia-Hwa
; APPLICANT: Simpson, David A.
; APPLICANT: Chang, Theresa Li-Yun
; APPLICANT: Xu, Qiang
; APPLICANT: Lewicki, John A.
; APPLICANT: Osel, Inc.
; TITLE OF INVENTION: Lactobacilli Expressing Biologically Active
; FILE REFERENCE: 016976-000610US
; CURRENT APPLICATION NUMBER: US/10/383,834
; PRIOR FILING DATE: 2003-06-03
; PRIOR APPLICATION NUMBER: US 60/362,945
; PRIOR FILING DATE: 2002-03-08
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: anchor sequence
US-10-383-834-4

Query Match 39.1%; Score 43; DB 6; Length 238;
Best Local Similarity 38.5%; Pred. No. 42;
Matches 10; Conservative 2; Mismatches 4; Indels 10; Gaps 1;

Qy 1 NNAGSV-----WTPKKPGLR 16
Db 31 SNAGYDPVTGKTTMNPWPAKQGLR 56

RESULT 8

US-60-494-568-17
; Sequence 17, Application US/60494568

; GENERAL INFORMATION:
; APPLICANT: ECOPIA BIOSCIENCES INC.
; APPLICANT: Zazopoulos, Emmanuel
; TITLE OF INVENTION: GENES AND PROTEINS FOR THE BIOSYNTHESIS OF POLYKETIDE COMPOUND
; FILE REFERENCE: 3004-2P
; CURRENT APPLICATION NUMBER: US/60/494,568
; CURRENT FILING DATE: 2003-08-13
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 17
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Streptomyces aizunensis
US-60-494-568-17

Query Match 39.1%; Score 43; DB 7; Length 367;
Best Local Similarity 50.0%; Pred. No. 63;
Matches 9; Conservative 4; Mismatches 3; Indels 2; Gaps 1;

Qy 2 NAGSV--EWTPKKPGLRR 17
Db 189 NSGAVLPDWLEKPGRRR 206

RESULT 9

US-10-374-979-108
; Sequence 108, Application US/10374979

; GENERAL INFORMATION:
; APPLICANT: John P. Carulli et al.
; TITLE OF INVENTION: THE HIGH BONE MASS GENE OF 11q13.3
; FILE REFERENCE: 032796-021
; CURRENT APPLICATION NUMBER: US/10/374,979
; CURRENT FILING DATE: 2003-03-04
; PRIOR APPLICATION NUMBER: US 09/544,398
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 09/543,771
; PRIOR FILING DATE: 2000-04-05

; PRIOR APPLICATION NUMBER: US 09/229,319
; PRIOR FILING DATE: 1999-01-13
; PRIOR APPLICATION NUMBER: US 60/071,449
; PRIOR FILING DATE: 1998-01-13
; PRIOR APPLICATION NUMBER: US 60/105,511
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 109
; SEQ ID NO 108
; LENGTH: 2861
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-374-979-108

Query Match 39.1%; Score 43; DB 6; Length 2861;
Best Local Similarity 53.8%; Pred. No. 4.5e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 8 WTPKKPGLRRROT 20
Db 2293 WSPQPRARQROT 2305

RESULT 10

US-10-331-496A-89
; Sequence 89, Application US/10331496A

; GENERAL INFORMATION:
; APPLICANT: FRANTZ, JRETCHEM
; APPLICANT: HILLAN, KENNETH J.
; APPLICANT: PHILLIPS, HEIDI S.
; APPLICANT: POLAKIS, PAUL
; APPLICANT: SMITH, VICTORIA
; APPLICANT: SPENCER, SUSAN D.
; APPLICANT: WILLIAMS, P. MICKEY
; APPLICANT: WU, THOMAS D.
; APPLICANT: ZHANG, ZEMIN
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; FILE REFERENCE: P5014R1-PCT
; CURRENT APPLICATION NUMBER: US/10/331,496A
; CURRENT FILING DATE: 2002-12-30
; PRIOR APPLICATION NUMBER: US 60/345,444
; PRIOR FILING DATE: 2002-01-02
; PRIOR APPLICATION NUMBER: US 60/351,885
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: US 60/360,066
; PRIOR FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: US 60/362,004
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/366,869
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: US 60/366,284
; PRIOR FILING DATE: 2002-03-21
; PRIOR APPLICATION NUMBER: US 60/368,679
; PRIOR FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: US 60/404,809
; PRIOR FILING DATE: 2002-08-19
; PRIOR APPLICATION NUMBER: US 60/405,645
; PRIOR FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 95
; SEQ ID NO 89
; LENGTH: 2861
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-331-496A-89

Query Match 39.1%; Score 43; DB 6; Length 2861;
Best Local Similarity 53.8%; Pred. No. 4.5e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 8 WTPKKPGLRRROT 20
Db 2293 WSPQPRARQROT 2305

```
RESULT 11
; Sequence 10183, Application US/10273573
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-066
; CURRENT APPLICATION NUMBER: US/10/273,573
; CURRENT FILING DATE: 2002-10-18
; PRIOR FILING DATE: 2002-04-18
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: 09/522,929
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 10994
; SOFTWARE: Custom
; SEQ ID NO 10183
; LENGTH: 57
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(57)
; OTHER INFORMATION: Xaa = x or * as defined in Table 2
US-10-273-573-10183

Query Match      38.2%; Score 42; DB 6; Length 57;
Best Local Similarity 60.0%; Pred. No. 15;
Matches 6; Conservative 3; Mismatches 1; Indels 1; Gaps 0;

QY      8 WTPKKPGLRR 17
      1:1: 111:1
DB      24 WSPPLGLKR 33

RESULT 12
PCT-US03-23249-34
; Sequence 34, Application PC/TUS0323249
; GENERAL INFORMATION:
; APPLICANT: ELLIOTT, Vicki S.; SWARNAKAR, Anita;
; APPLICANT: GRIFFIN, Jennifer A.; LEE, Ernestine A.;
; APPLICANT: SPRAGUE, William W.; HAFALIA, April J.A.;
; APPLICANT: LEE, Soo Yeun; KABLE, Amy E.;
; APPLICANT: ISON, Craig H.; KHARE, Reena;
; APPLICANT: CHAWLA, Narinder K.; MARQUIS, Joseph P.;
; APPLICANT: JIANG Xin; JACKSON, Alan A.;
; APPLICANT: BECHA, Shanya D.; EMERLING, Brooke M.;
; APPLICANT: JIN, Fei; WILSON, Amy D.;
; APPLICANT: RICHARDSON, Thomas W.; YANG, Junming;
; APPLICANT: BAUGHN, Mariah R.; GANDHI, Ameena R.;
; APPLICANT: NGUYEN, Dannel B.; RAMKUMAR, Jayalaxmi;
; APPLICANT: KALLICK, Deborah A.; KEARNEY, Liam;
; APPLICANT: LU, Dyung Aina M.; GIETZEN, Kimberly J.;
; APPLICANT: TRIBOULEY, Catherine M.; LAL, Preeti G.;
; APPLICANT: BLAKE, Julie J.; LU, Yan;
; APPLICANT: ARVIZU, Chandra S.
; TITLE OF INVENTION: PROTEIN MODIFICATION AND MAINTENANCE MOLECULES
; FILE REFERENCE: PF-1478 PCT
; CURRENT APPLICATION NUMBER: PCT/US03/23249
; CURRENT FILING DATE: 2003-07-23
; PRIOR APPLICATION NUMBER: US 60/398,143
; PRIOR FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: US 60/402,458
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/403,289
; PRIOR FILING DATE: 2002-08-12
; PRIOR APPLICATION NUMBER: US 60/406,472
; PRIOR FILING DATE: 2002-08-27
; PRIOR APPLICATION NUMBER: US 60/409,354
; PRIOR FILING DATE: 2002-09-06
; NUMBER OF SEQ ID NOS: 124
; SOFTWARE: PERL Program

Query Match      37.7%; Score 41.5; DB 6; Length 425;
Best Local Similarity 45.5%; Pred. No. 1.2e+02;
Matches 10; Conservative 1; Mismatches 6; Indels 5; Gaps 1;
```

```
; SEQ ID NO 34
; LENGTH: 271
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No: 7518798CD1
PCT-US03-23249-34

Query Match      37.7%; Score 41.5; DB 1; Length 271;
Best Local Similarity 45.5%; Pred. No. 77;
Matches 10; Conservative 1; Mismatches 6; Indels 5; Gaps 1;

QY      4 GSVW-----TPKKPGLRRRT 20
      1:1: 1111:1
DB      238 GIVSWGSCALPKPKGVYTRVT 259

RESULT 13
US-10-467-042-8
; Sequence 8, Application US/10467042
; GENERAL INFORMATION:
; APPLICANT: WARREN, Bridget A.; HONCHELL, Cynthia D.;
; APPLICANT: LU, Yan; CHAWLA, Narinder K.;
; APPLICANT: BURFORD, Neil; DELEGEANE, Angelo M.;
; APPLICANT: GANDHI, Ameena R.; BAUGHN, Mariah R.;
; APPLICANT: GRIFFIN, Jennifer A.; GIETZEN, Kimberly J.;
; APPLICANT: LU, Dyung Aina M.; ISON, Craig H.;
; APPLICANT: RAMKUMAR, Jayalaxmi; TANG, Y. Tom;
; APPLICANT: LAL, Preeti G.; BOROWSKI, Mark L.;
; APPLICANT: DUGGAN, Brendan M.; HAFALIA, April J.A.;
; APPLICANT: ARVIZU, Chandra S.; THANGAVELU, Kavitha;
; APPLICANT: YAO, Monique G.; ELLIOTT, Vicki S.;
; APPLICANT: DING, Li; YUE, Henry;
; APPLICANT: LEE, Sally; SWARNAKAR, Anita;
; APPLICANT: TRAN, Uyen K.; XU, Yumling
; TITLE OF INVENTION: PROTEIN MODIFICATION AND MAINTENANCE MOLECULES
; FILE REFERENCE: PI-0361 USN
; CURRENT APPLICATION NUMBER: US/10/467,042
; CURRENT FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: PCT/US02/02813
; PRIOR FILING DATE: 2002-01-30
; PRIOR APPLICATION NUMBER: US 60/265,705
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: US 60/266,762
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: US 60/269,581
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 60/271,198
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: US 60/272,813
; PRIOR FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: US 60/275,586
; PRIOR FILING DATE: 2001-03-13
; PRIOR APPLICATION NUMBER: US 60/278,505
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 60/280,539
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PERL Program
; SEQ ID NO 8
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No: 7481056CD1
US-10-467-042-8

Query Match      37.7%; Score 41.5; DB 6; Length 425;
Best Local Similarity 45.5%; Pred. No. 1.2e+02;
Matches 10; Conservative 1; Mismatches 6; Indels 5; Gaps 1;
```

Qy 4 GSVEN-----TPKKPGLRRROT 20
| | | | | | | | | | | | | | | | | |
Db 392 GIVSWGOSCALPKKPGVYTRVT 413

RESULT 14

US-10-308-817-103
; Sequence 103, Application US/10308817
; GENERAL INFORMATION:
; APPLICANT: Rother, Russell
; TITLE OF INVENTION: HYBRID ANTIBODIES
; FILE REFERENCE: 1087-37
; CURRENT APPLICATION NUMBER: US/10/308,817
; CURRENT FILING DATE: 2002-12-03
; NUMBER OF SEQ ID NOS: 195
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 103
; LENGTH: 95
; TYPE: PRT
; ORGANISM: human
US-10-308-817-103

Query Match 37.3%; Score 41; DB 6; Length 95;
Best Local Similarity 56.2%; Pred. No. 33;
Matches 9; Conservative 1; Mismatches 4; Indels 2; Gaps 1;

Qy 1 NNAGS--SVETPKKPG 14
| | | | | | | | | | | | | | | | | |
Db 25 NNIGSKNVHWYQQRPG 40

RESULT 15

US-10-308-817-107
; Sequence 107, Application US/10308817
; GENERAL INFORMATION:
; APPLICANT: Rother, Russell
; TITLE OF INVENTION: HYBRID ANTIBODIES
; FILE REFERENCE: 1087-37
; CURRENT APPLICATION NUMBER: US/10/308,817
; CURRENT FILING DATE: 2002-12-03
; NUMBER OF SEQ ID NOS: 195
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 107
; LENGTH: 96
; TYPE: PRT
; ORGANISM: human
US-10-308-817-107

Query Match 37.3%; Score 41; DB 6; Length 96;
Best Local Similarity 56.2%; Pred. No. 33;
Matches 9; Conservative 1; Mismatches 4; Indels 2; Gaps 1;

Qy 1 NNAG--SVETPKKPG 14
| | | | | | | | | | | | | | | | | |
Db 25 NNIGSKNVHWYQQRPG 40

Search completed: August 27, 2003, 09:56:01
Job time : 22 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2003, 09:41:14 ; Search time 348 seconds
(without alignments)
50.021 Million cell updates/sec

Title: US-10-042-417A-91

Perfect score: 110
Sequence: 1 NNAGSVEWTPKKPLRRRQT 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 5580241 seqs, 870357830 residues

Total number of hits satisfying chosen parameters: 5580241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents_AA_Main:*

1:	/cgn2_6/ptodata/2/paa/PCTUS_COMB.pep.*
2:	/cgn2_6/ptodata/2/paa/US06_COMB.pep.*
3:	/cgn2_6/ptodata/2/paa/US07_COMB.pep.*
4:	/cgn2_6/ptodata/2/paa/US080_COMB.pep.*
5:	/cgn2_6/ptodata/2/paa/US081_COMB.pep.*
6:	/cgn2_6/ptodata/2/paa/US082_COMB.pep.*
7:	/cgn2_6/ptodata/2/paa/US083_COMB.pep.*
8:	/cgn2_6/ptodata/2/paa/US084_COMB.pep.*
9:	/cgn2_6/ptodata/2/paa/US085_COMB.pep.*
10:	/cgn2_6/ptodata/2/paa/US086_COMB.pep.*
11:	/cgn2_6/ptodata/2/paa/US087_COMB.pep.*
12:	/cgn2_6/ptodata/2/paa/US088_COMB.pep.*
13:	/cgn2_6/ptodata/2/paa/US089_COMB.pep.*
14:	/cgn2_6/ptodata/2/paa/US090_COMB.pep.*
15:	/cgn2_6/ptodata/2/paa/US091_COMB.pep.*
16:	/cgn2_6/ptodata/2/paa/US092_COMB.pep.*
17:	/cgn2_6/ptodata/2/paa/US093_COMB.pep.*
18:	/cgn2_6/ptodata/2/paa/US094_COMB.pep.*
19:	/cgn2_6/ptodata/2/paa/US095_COMB.pep.*
20:	/cgn2_6/ptodata/2/paa/US096_COMB.pep.*
21:	/cgn2_6/ptodata/2/paa/US097A_COMB.pep.*
22:	/cgn2_6/ptodata/2/paa/US097B_COMB.pep.*
23:	/cgn2_6/ptodata/2/paa/US098_COMB.pep.*
24:	/cgn2_6/ptodata/2/paa/US099A_COMB.pep.*
25:	/cgn2_6/ptodata/2/paa/US099B_COMB.pep.*
26:	/cgn2_6/ptodata/2/paa/US100_COMB.pep.*
27:	/cgn2_6/ptodata/2/paa/US101_COMB.pep.*
28:	/cgn2_6/ptodata/2/paa/US102_COMB.pep.*
29:	/cgn2_6/ptodata/2/paa/US103_COMB.pep.*
30:	/cgn2_6/ptodata/2/paa/US104_COMB.pep.*
31:	/cgn2_6/ptodata/2/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	91	82.7	19	1	PCT-US02-00311-85
2	91	82.7	19	1	PCT-US99-19560-85
					Sequence 85, Appl
					Sequence 85, Appl

3	91	82.7	19	17	US-09-385-219-85	Sequence 85, Appl
4	91	82.7	19	17	US-09-385-219A-85	Sequence 85, Appl
5	91	82.7	19	26	US-10-042-417-85	Sequence 85, Appl
6	91	82.7	25	19	US-09-587-473-1	Sequence 1, Appl
7	91	82.7	87	21	US-09-724-676-67528	Sequence 67528, A
8	91	82.7	87	21	US-09-724-676A-67528	Sequence 3, Appl
9	91	82.7	198	6	US-08-275-983A-3	Sequence 2, Appl
10	91	82.7	198	11	US-08-765-702-2	Sequence 2, Appl
11	91	82.7	198	11	US-08-765-702A-2	Sequence 2, Appl
12	91	82.7	198	12	US-08-893-276A-2	Sequence 2, Appl
13	91	82.7	198	12	US-08-896-920-2	Sequence 2, Appl
14	91	82.7	198	12	US-08-897-333-2	Sequence 2, Appl
15	91	82.7	198	14	US-09-099-857-2	Sequence 2, Appl
16	91	82.7	198	17	US-09-378-517-6	Sequence 6, Appl
17	91	82.7	198	18	US-09-483-597-6	Sequence 6, Appl
18	91	82.7	198	20	US-09-614-099-9	Sequence 9, Appl
19	91	82.7	198	21	US-09-719-755-2	Sequence 2, Appl
20	91	82.7	198	23	US-09-865-018-2	Sequence 2, Appl
21	91	82.7	198	23	US-09-865-018A-6	Sequence 6, Appl
22	91	82.7	198	23	US-09-865-018B-2	Sequence 2, Appl
23	91	82.7	198	25	US-09-970-561-2	Sequence 2, Appl
24	91	82.7	198	26	US-10-087-192-1164	Sequence 1164, Ap
25	91	82.7	198	27	US-10-170-385-303	Sequence 303, App
26	91	82.7	198	29	US-10-302-812-78	Sequence 78, Appl
27	91	82.7	198	30	US-10-458-108-9	Sequence 9, Appl
28	91	82.7	198	31	US-60-443-566-2951	Sequence 2951, Ap
29	91	82.7	198	31	US-60-452-680-21593	Sequence 21593, A
30	91	82.7	198	31	US-60-455-444-7024	Sequence 7024, Ap
31	91	82.7	198	31	US-60-465-241-7024	Sequence 7024, Ap
32	91	82.7	198	31	US-60-470-166-1212	Sequence 1212, Ap
33	91	82.7	218	6	US-08-275-983-3	Sequence 3, Appl
34	91	82.7	365	21	US-09-718-233-8	Sequence 8, Appl
35	91	82.7	380	21	US-09-718-233-6	Sequence 6, Appl
36	91	82.7	391	21	US-09-718-233-2	Sequence 2, Appl
37	74	67.3	38	6	US-08-275-983-2	Sequence 2, Appl
38	74	67.3	183	16	US-09-215-224-42	Sequence 42, Appl
39	74	67.3	183	16	US-09-215-224-47	Sequence 47, Appl
40	74	67.3	183	16	US-09-215-224D-42	Sequence 42, Appl
41	74	67.3	183	16	US-09-215-224D-47	Sequence 47, Appl
42	74	67.3	197	6	US-08-275-983A-2	Sequence 2, Appl
43	74	67.3	197	11	US-08-765-702-4	Sequence 4, Appl
44	74	67.3	197	11	US-08-765-702B-4	Sequence 4, Appl
45	74	67.3	197	12	US-08-896-920-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1
PCT-US02-00311-85
; Sequence 85, Application PC/TUS0200311
; GENERAL INFORMATION:
; APPLICANT: Pagano, M.
; TITLE OF INVENTION: METHODS TO IDENTIFY COMPOUNDS USEFUL FOR THE TREATMENT OF
; PROLIFERATIVE AND DIFFERENTIATIVE DISORDERS
; FILE OF INVENTION: 5914-090-228
; CURRENT APPLICATION NUMBER: PCT/US02/00311
; CURRENT FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: 60/260,179
; PRIOR FILING DATE: 2001-01-05
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 85
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Homo sapiens
; PCT-US02-00311-85

Query Match 82.7%; Score 91; DB 1; Length 19;
Best Local Similarity 94.7%; Pred No. 2.4e-05;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 NAGSVEWTPKKPLRRRQT 20

Db 1 NAGSVEQTPKKPGLRRRQT 19
||||| |||||||

RESULT 2

PCT-US99-19560-85
; Sequence 85, Application PC/TUS9919560
; GENERAL INFORMATION:
; APPLICANT: NEW YORK UNIVERSITY
; TITLE OF INVENTION: NOVEL UBIQUITIN LIGASES AS THERAPEUTIC TARGETS
; FILE REFERENCE: 5914-081-228
; CURRENT APPLICATION NUMBER: PCT/US99/19560
; CURRENT FILING DATE: 1999-08-31
; EARLIER APPLICATION NUMBER: 60/098,355
; EARLIER FILING DATE: 1998-08-28
; EARLIER APPLICATION NUMBER: 60/118,568
; EARLIER FILING DATE: 1999-02-03
; EARLIER APPLICATION NUMBER: 60/124,449
; EARLIER FILING DATE: 1999-03-15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 85
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US99-19560-85

Query Match 82.7%; Score 91; DB 1; Length 19;
Best Local Similarity 94.7%; Pred. No. 2.4e-05;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 NAGSVEWTPKKPGLRRRQT 20
||||| |||||||
Db 1 NAGSVEQTPKKPGLRRRQT 19

RESULT 3

US-09-385-219-85
; Sequence 85, Application US/09385219
; GENERAL INFORMATION:
; APPLICANT: Chlaaur, D.
; APPLICANT: Pagano, M.
; APPLICANT: Latres, E.
; TITLE OF INVENTION: NOVEL UBIQUITIN LIGASES AS THERAPEUTIC TARGETS
; FILE REFERENCE: 5914-081
; CURRENT APPLICATION NUMBER: US/09/385,219
; CURRENT FILING DATE: 1999-08-27
; EARLIER APPLICATION NUMBER: 60/098,355
; EARLIER FILING DATE: 1998-08-28
; EARLIER APPLICATION NUMBER: 60/118,568
; EARLIER FILING DATE: 1999-02-03
; EARLIER APPLICATION NUMBER: 60/124,449
; EARLIER FILING DATE: 1999-03-15
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 85
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-385-219-85

Query Match 82.7%; Score 91; DB 17; Length 19;
Best Local Similarity 94.7%; Pred. No. 2.4e-05;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 NAGSVEWTPKKPGLRRRQT 20
||||| |||||||
Db 1 NAGSVEQTPKKPGLRRRQT 19

RESULT 4

US-09-385-219A-85
; Sequence 85, Application US/09385219A

; GENERAL INFORMATION:
; APPLICANT: Chlaaur, D.
; APPLICANT: Pagano, M.
; APPLICANT: Latres, E.
; TITLE OF INVENTION: NOVEL UBIQUITIN LIGASES AS THERAPEUTIC TARGETS
; FILE REFERENCE: 5914-081
; CURRENT APPLICATION NUMBER: US/09/385,219A
; CURRENT FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: 60/098,355
; PRIOR FILING DATE: 1998-08-28
; PRIOR APPLICATION NUMBER: 60/118,568
; PRIOR FILING DATE: 1999-02-03
; PRIOR APPLICATION NUMBER: 60/124,449
; PRIOR FILING DATE: 1999-03-15
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 85
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: Phosphorylation
; LOCATION: 8
; OTHER INFORMATION: Phosothreonine
US-09-385-219A-85

Query Match 82.7%; Score 91; DB 17; Length 19;
Best Local Similarity 94.7%; Pred. No. 2.4e-05;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 NAGSVEWTPKKPGLRRRQT 20
||||| |||||||
Db 1 NAGSVEQTPKKPGLRRRQT 19

RESULT 5

US-10-042-417-85
; Sequence 85, Application US/10042417
; GENERAL INFORMATION:
; APPLICANT: Pagano, M.
; TITLE OF INVENTION: METHODS TO IDENTIFY COMPOUNDS USEFUL FOR THE TREATMENT OF
; FILE REFERENCE: 5914-090-999
; CURRENT APPLICATION NUMBER: US/10/042,417
; CURRENT FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: 60/260,179
; PRIOR FILING DATE: 2001-01-5
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 85
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-042-417-85

Query Match 82.7%; Score 91; DB 26; Length 19;
Best Local Similarity 94.7%; Pred. No. 2.4e-05;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 NAGSVEWTPKKPGLRRRQT 20
||||| |||||||
Db 1 NAGSVEQTPKKPGLRRRQT 19

RESULT 6

US-09-587-473-1
; Sequence 1, Application US/09587473
; GENERAL INFORMATION:
; APPLICANT: Zhang, Hui
; APPLICANT: Tsvetkov, Lyuben
; TITLE OF INVENTION: Protein Knockout Technology
; FILE REFERENCE: 44574-5047-WO
; CURRENT APPLICATION NUMBER: US/09/587,473

; CURRENT FILING DATE: 2000-06-05
; PRIOR APPLICATION NUMBER: US 60/137,494
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (14)
; OTHER INFORMATION: PHOSPHORYLATION- Peptide is carboxyl end of p27.
US-09-587-473-1

Query Match 82.7%; Score 91; DB 19; Length 25;
Best Local Similarity 94.7%; Pred. No. 3.2e-05;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 NAGSVEQTPKKPGLRRROT 20
||||| |||||||||
DB 7 NAGSVEQTPKKPGLRRROT 25

RESULT 7
US-09-724-676-67528
; Sequence 67528, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 67528
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-67528

Query Match 82.7%; Score 91; DB 21; Length 87;
Best Local Similarity 94.7%; Pred. No. 0.00011;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 NAGSVEQTPKKPGLRRROT 20
||||| |||||||||
DB 69 NAGSVEQTPKKPGLRRROT 87

RESULT 8
US-09-724-676A-67528
; Sequence 67528, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 67528
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-67528

Query Match 82.7%; Score 91; DB 21; Length 87;
Best Local Similarity 94.7%; Pred. No. 0.00011;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 NAGSVEQTPKKPGLRRROT 20
||||| |||||||||

DB 69 NAGSVEQTPKKPGLRRROT 87

RESULT 9
US-08-275-983A-3
; Sequence 3, Application US/08275983A
; GENERAL INFORMATION:
; APPLICANT: Massague, Joan
; APPLICANT: Roberts, James M.
; APPLICANT: Koff, Andrew
; APPLICANT: Polyak, Kornelia
; TITLE OF INVENTION: ISOLATED p27 PROTEIN, NUCLEIC ACID
; TITLE OF INVENTION: MOLECULES ENCODING SAME, METHODS OF
; TITLE OF INVENTION: IDENTIFYING AGENTS ACTING ON SAME, AND
; TITLE OF INVENTION: USES OF SAID AGENTS
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/275,983A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 1747/44652-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 198 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-275-983A-3

Query Match 82.7%; Score 91; DB 6; Length 198;
Best Local Similarity 94.7%; Pred. No. 0.00026;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 NAGSVEQTPKKPGLRRROT 20
||||| |||||||||
DB 180 NAGSVEQTPKKPGLRRROT 198

RESULT 10
US-08-765-702-2
; Sequence 2, Application US/08765702
; GENERAL INFORMATION:
; APPLICANT: Massague, Joan
; APPLICANT: Roberts, James M.
; APPLICANT: Koff, Andrew
; TITLE OF INVENTION: ISOLATED p27 PROTEIN, NUCLEIC ACID
; TITLE OF INVENTION: MOLECULES ENCODING SAME, METHODS OF
; TITLE OF INVENTION: WHICH ACT ON p27 PROTEIN, AND USES OF SAID AGENTS
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLHY, HOAG & ELIOT LLP
; STREET: One Post Office Square

```

; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/765,702
; FILING DATE: 28-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MIV-079.04
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 198 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-765-702-2

```

```

Query Match      82.7%; Score 91; DB 11; Length 198;
Best Local Similarity 94.7%; Pred. No. 0.00026;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 NAGSVEWTPKPKGLRRROT 20
      ||||| ||||| |||||
Db      180 NAGSVEQTPKPKGLRRROT 198

```

```

RESULT 11
US-08-765-702B-2
; Sequence 2, Application US/08765702B
; GENERAL INFORMATION:
; APPLICANT: Massague, Joan
;               Roberts, James M.
;               Polyak, Kornelia
; TITLE OF INVENTION: ISOLATED p27 PROTEIN, NUCLEIC ACID
;               MOLECULES ENCODING SAME, METHODS OF IDENTIFYING AGENTS
;               ACTING ON SAME, AND USES OF SAID AGENTS
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELLIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPad
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/765,702B
; FILING DATE: 28-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MIV-079.04
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 198 amino acids

```

```

; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-08-765-702B-2

```

```

Query Match      82.7%; Score 91; DB 11; Length 198;
Best Local Similarity 94.7%; Pred. No. 0.00026;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 NAGSVEWTPKPKGLRRROT 20
      ||||| ||||| |||||
Db      180 NAGSVEQTPKPKGLRRROT 198

```

```

RESULT 12
US-08-893-276A-2
; Sequence 2, Application US/08893276A
; GENERAL INFORMATION:
; APPLICANT: Michele Pagano
; APPLICANT: Giulio Draetta
; APPLICANT: Mark Rolfe
; APPLICANT: Massimo Loda
; TITLE OF INVENTION: Reagents and Methods for Diagnosis and Prognosis
;               of Proliferative Disorders
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley, Hoag & Elliot, LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII (text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/893,276A
; FILING DATE: 15-JULY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: Provisional No. 60/036,690
; FILING DATE: 31-JANUARY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MIV-093.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 832-1000
; TELEFAX: (617) 832-7000
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 198 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-893-276A-2

```

```

Query Match      82.7%; Score 91; DB 12; Length 198;
Best Local Similarity 94.7%; Pred. No. 0.00026;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 NAGSVEWTPKPKGLRRROT 20
      ||||| ||||| |||||
Db      180 NAGSVEQTPKPKGLRRROT 198

```

```

RESULT 13
US-08-896-920-2
; Sequence 2, Application US/08896920
; GENERAL INFORMATION:
; APPLICANT: Roberts, James M.

```



```
; APPLICANT: Massague, Joan
; APPLICANT: Koff, Andrew
; APPLICANT: Polyak, Kornelia
; TITLE OF INVENTION: ISOLATED P27 PROTEIN AND METHODS FOR ITS
; TITLE OF INVENTION: PRODUCTION AND USE
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/896,920
; FILING DATE: 18-JUL-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MIV-079.07
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 198 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-896-920-2
```

```
Query Match      82.7%; Score 91; DB 12; Length 198;
Best Local Similarity 94.7%; Pred. No. 0.00026;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy      2 NAGSVETPKPKGLRRRT 20
      ||||| ||||| |||||
Db      180 NAGSVEQTPKKPGLRRRT 198
```

```
RESULT 14
US-08-897-333-2
; Sequence 2, Application US/08897333
; GENERAL INFORMATION:
; APPLICANT: Nabel, Gary J.
; APPLICANT: Nabel, Elizabeth G.
; TITLE OF INVENTION: METHOD FOR TREATING VASCULAR PROLIFERATIVE DISEASES
; TITLE OF INVENTION: WITH P27 AND FUSIONS THEREOF
; FILE REFERENCE: 8642/4
; CURRENT APPLICATION NUMBER: US/08/897,333
; CURRENT FILING DATE: 1997-07-21
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-897-333-2
```

```
Query Match      82.7%; Score 91; DB 12; Length 198;
Best Local Similarity 94.7%; Pred. No. 0.00026;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy      2 NAGSVETPKPKGLRRRT 20
      ||||| ||||| |||||
Db      180 NAGSVEQTPKKPGLRRRT 198
```

```
RESULT 15
US-09-099-857-2
; Sequence 2, Application US/09099857
; GENERAL INFORMATION:
; APPLICANT: Nandabalan, Krishnan
; APPLICANT: Yang, MeiJa
; TITLE OF INVENTION: P27(Kip-1)-FKBP-12 Protein Complexes
; FILE REFERENCE: Cura-14 US: P27(Kip-1)-FKBP-12 Complex
; CURRENT APPLICATION NUMBER: US/09/099,857
; CURRENT FILING DATE: 1998-06-18
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-099-857-2
```

```
Query Match      82.7%; Score 91; DB 14; Length 198;
Best Local Similarity 94.7%; Pred. No. 0.00026;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy      2 NAGSVETPKPKGLRRRT 20
      ||||| ||||| |||||
Db      180 NAGSVEQTPKKPGLRRRT 198
```

```
Search completed: August 27, 2003, 09:55:26
Job time : 350 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 27, 2003, 09:45:48 ; Search time 283 Seconds
(without alignments)

9.667 Million cell updates/sec

Title: US-10-042-417A-91

Perfect score: 110

Sequence: 1 NNAGSVEWTKKPKGLRRROT 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 510680 seqs, 136781880 residues

Total number of hits satisfying chosen parameters: 510680

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA:*

```
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	91	82.7	19	14	US-10-042-417-85
2	91	82.7	198	11	US-09-865-018-2
3	91	82.7	198	11	US-09-970-561-2
4	91	82.7	365	8	US-08-902-572-8
5	91	82.7	380	8	US-08-902-572-6
6	91	82.7	391	8	US-08-902-572-2
7	74	67.3	197	10	US-09-865-018-4
8	45	40.9	355	15	US-10-128-714-3463
9	45	40.9	355	15	US-10-128-714-8463
10	45	40.9	518	10	US-09-803-589-10
11	45	40.9	551	12	US-10-365-227-16
12	45	40.9	551	14	US-10-105-929-16
13	45	40.9	727	10	US-09-445-023A-12
14	45	40.9	727	15	US-10-097-597-12
15	45	40.9	727	15	US-10-097-580-12

```
Sequence 52, Appl
Sequence 52, Appl
Sequence 4, Appl
Sequence 7, Appl
Sequence 41468, A
Sequence 4, Appl
Sequence 512, App
Sequence 10110, A
Sequence 6, Appl
Sequence 2, Appl
Sequence 1748, Ap
Sequence 19, Appl
Sequence 764, App
Sequence 9926, Ap
Sequence 18, Appl
Sequence 42, Appl
Sequence 35, Appl
Sequence 92, Appl
Sequence 16, Appl
Sequence 4213, Ap
Sequence 535, App
Sequence 430, App
Sequence 1830, Ap
Sequence 10192, A
Sequence 12, Appl
Sequence 20, Appl
```

ALIGNMENTS

RESULT 1

```
US-10-042-417-85
; Sequence 85, Application US/10042417
; Publication No. US20020123082A1
; GENERAL INFORMATION:
; APPLICANT: Pagano, M.
; TITLE OF INVENTION: METHODS TO IDENTIFY COMPOUNDS USEFUL FOR THE TREATMENT OF
; PROLIFERATIVE AND DIFFERENTIATIVE DISORDERS
; FILE REFERENCE: 5914-090-999
; CURRENT APPLICATION NUMBER: US/10/042,417
; CURRENT FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: 60/260,179
; PRIOR FILING DATE: 2001-01-5
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 85
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-042-417-85
```

```
Query Match      82.7%  Score 91:  DB 14:  Length 19;
Best Local Similarity 94.7%;  Pred No. 1.8e-06;
Matches 18;  Conservative 0;  Mismatches 1;  Indels 0;  Gaps 0;
```

```
Oy  2  NAGSVEWTKKPKGLRRROT 20
      ||||| ||||| |||||
Db   1  NAGSVEQTEKKPKGLRRROT 19
```

RESULT 2

```
US-09-865-018-2
; Sequence 2, Application US/09865018
; Patent No. US20020110886A1
; GENERAL INFORMATION:
; APPLICANT: Massague, Joan
; Roberts, James M.
; Koff, Andrew
```

Polvak, Kornelia
TITLE OF INVENTION: ISOLATED p27 PROTEIN AND METHOD FOR ITS PRODUCTION AND USE
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HOAG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/865.018
FILING DATE: 24-May-2001
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/854.039
FILING DATE: 09-MAY-1997
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: MIV-079.04
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-1000
TELEFAX: 617-832-7000
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 198 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-865-018-2

Query Match 82.7%; Score 91; DB 10; Length 198;
Best Local Similarity 94.7%; Pred. No. 1.7e-05;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 NAGSVEWTPKKPGLRRRQT 20
||||| |||||||||
DB 180 NAGSVEQTPKKPGLRRRQT 198

RESULT 3
US-09-970-561-2
Sequence 2, Application US/09970561
Publication No. US20030023034A1
GENERAL INFORMATION:
APPLICANT: Nandabalan, Krishnan
APPLICANT: Yang, Melja
TITLE OF INVENTION: p27(kip-1)-FKBP-12 Protein Complexes
FILE REFERENCE: Cura-14 CON (15966-514 CON)
CURRENT APPLICATION NUMBER: US/09/970.561
CURRENT FILING DATE: 2001-10-03
PRIOR APPLICATION NUMBER: PCT/US99/13659
PRIOR FILING DATE: 1999-06-18
PRIOR APPLICATION NUMBER: USSN 09/719,755
PRIOR FILING DATE: 2000-12-15
NUMBER OF SEQ ID NOS: 6
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 2
LENGTH: 198
TYPE: PRT
ORGANISM: Homo sapiens
US-09-970-561-2

Query Match 82.7%; Score 91; DB 11; Length 198;
Best Local Similarity 94.7%; Pred. No. 1.7e-05;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 NAGSVEWTPKKPGLRRRQT 20
||||| |||||||||
DB 180 NAGSVEQTPKKPGLRRRQT 198
RESULT 4
US-08-902-572-8
Sequence 8, Application US/08902572
Publication No. US20020068706A1
GENERAL INFORMATION:
APPLICANT: Gyuris, Jenő
APPLICANT: Lamphere, Lou
APPLICANT: Beach, David H.
TITLE OF INVENTION: INHIBITORS OF CELL-CYCLE PROGRESSION AND
TITLE OF INVENTION: RELATED THERETO
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HOAG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/902.572
FILING DATE: 29-JUL-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: MIV-069.03
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-1000
TELEFAX: 617-832-7000
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 365 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-902-572-8

Query Match 82.7%; Score 91; DB 8; Length 365;
Best Local Similarity 94.7%; Pred. No. 3e-05;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 NAGSVEWTPKKPGLRRRQT 20
||||| |||||||||
DB 347 NAGSVEQTPKKPGLRRRQT 365

RESULT 5
US-08-902-572-6
Sequence 6, Application US/08902572
Publication No. US20020068706A1
GENERAL INFORMATION:
APPLICANT: Gyuris, Jenő
APPLICANT: Lamphere, Lou
APPLICANT: Beach, David H.
TITLE OF INVENTION: INHIBITORS OF CELL-CYCLE PROGRESSION AND
TITLE OF INVENTION: RELATED THERETO
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HOAG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA

```
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/902.572
; FILING DATE: 29-JUL-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MIV-069.03
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 380 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-902-572-6

Query Match      82.7%; Score 91; DB 8; Length 380;
Best Local Similarity 94.7%; Pred. No. 3.1e-05;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 NAGSVEWTPKKPGLRRRT 20
Db 362 NAGSVEQTPKKPGLRRRT 380

RESULT 6
US-08-902-572-2
; Sequence 2, Application US/08902572
; Publication No. US20020068706A1
; GENERAL INFORMATION:
; APPLICANT: Gyuris, Jeno
; APPLICANT: Lamphere, Lou
; APPLICANT: Beach, David H.
; TITLE OF INVENTION: INHIBITORS OF CELL-CYCLE PROGRESSION AND
; TITLE OF INVENTION: RELATED THERETO
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/902.572
; FILING DATE: 29-JUL-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MIV-069.03
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 391 amino acids
; TYPE: amino acid
```

```
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-902-572-2

Query Match      82.7%; Score 91; DB 8; Length 391;
Best Local Similarity 94.7%; Pred. No. 3.2e-05;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 NAGSVEWTPKKPGLRRRT 20
Db 186 NAGSVEQTPKKPGLRRRT 204

RESULT 7
US-09-865-018-4
; Sequence 4, Application US/09865018
; Patent No. US20020110886A1
; GENERAL INFORMATION:
; APPLICANT: Massague, Joan
; APPLICANT: Roberts, James M.
; APPLICANT: Koff, Andrew
; APPLICANT: Polyak, Kornelia
; TITLE OF INVENTION: ISOLATED p27 PROTEIN AND METHOD FOR ITS
; PRODUCTION AND USE
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/865.018
; FILING DATE: 24-May-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/854.039
; FILING DATE: 09-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MIV-079.04
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 197 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
; US-09-865-018-4

Query Match      67.3%; Score 74; DB 10; Length 197;
Best Local Similarity 82.4%; Pred. No. 0.0044;
Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 NAGSVEWTPKKPGLRRR 18
Db 180 NAGTVEQTPKKPGLRRQ 196

RESULT 8
US-10-128-714-3463
; Sequence 3463, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
```

APPLICANT: Jiang, Bo
APPLICANT: Hu, Wenqi
APPLICANT: Tishkoff, Daniel
APPLICANT: Zamudio, Carlos
APPLICANT: Eroszkin, Alexey M
APPLICANT: Lemieux, Sebastien M
TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
TITLE OF INVENTION: Methods of Use
FILE REFERENCE: 10182-018-999
CURRENT APPLICATION NUMBER: US/10/128,714
CURRENT FILING DATE: 2002-04-23
PRIOR APPLICATION NUMBER: US 60/285,697
PRIOR FILING DATE: 2001-04-23
PRIOR APPLICATION NUMBER: US 60/287,066
PRIOR FILING DATE: 2001-04-27
PRIOR APPLICATION NUMBER: US 60/295,890
PRIOR FILING DATE: 2001-06-05
PRIOR APPLICATION NUMBER: US 60/303,899
PRIOR FILING DATE: 2001-07-09
PRIOR APPLICATION NUMBER: US 60/316,362
PRIOR FILING DATE: 2001-08-31
NUMBER OF SEQ ID NOS: 8603
SOFTWARE: PatentIn version 3.1
SEQ ID NO 3463
LENGTH: 355
TYPE: PRT
ORGANISM: Aspergillus fumigatus
US-10-128-714-3463

Query Match 40.9%; Score 45; DB 15; Length 355;
Best Local Similarity 52.9%; Pred. No. 1.1e+02;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 2 NAGSVEWTPKKPGLRRR 18
Db 198 NAGKREGWPEAPGLNKR 214

RESULT 9
US-10-128-714-8463
Sequence 8463, Application US/10128714
Publication No. US20030119013A1
GENERAL INFORMATION:
APPLICANT: Jiang, Bo
APPLICANT: Hu, Wenqi
APPLICANT: Tishkoff, Daniel
APPLICANT: Zamudio, Carlos
APPLICANT: Eroszkin, Alexey M
APPLICANT: Lemieux, Sebastien M
TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
TITLE OF INVENTION: Methods of Use
FILE REFERENCE: 10182-018-999
CURRENT APPLICATION NUMBER: US/10/128,714
CURRENT FILING DATE: 2002-04-23
PRIOR APPLICATION NUMBER: US 60/285,697
PRIOR FILING DATE: 2001-04-23
PRIOR APPLICATION NUMBER: US 60/287,066
PRIOR FILING DATE: 2001-04-27
PRIOR APPLICATION NUMBER: US 60/295,890
PRIOR FILING DATE: 2001-06-05
PRIOR APPLICATION NUMBER: US 60/303,899
PRIOR FILING DATE: 2001-07-09
PRIOR APPLICATION NUMBER: US 60/316,362
PRIOR FILING DATE: 2001-08-31
NUMBER OF SEQ ID NOS: 8603
SOFTWARE: PatentIn version 3.1
SEQ ID NO 8463
LENGTH: 355
TYPE: PRT
ORGANISM: Aspergillus fumigatus
US-10-128-714-8463

Query Match 40.9%; Score 45; DB 15; Length 355;

Best Local Similarity 52.9%; Pred. No. 1.1e+02;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 2 NAGSVEWTPKKPGLRRR 18
Db 198 NAGKREGWPEAPGLNKR 214

RESULT 10
US-09-803-589-10
Sequence 10, Application US/09803589
Patent No. US20020112251A1
GENERAL INFORMATION:
APPLICANT: McCarthy, Sean A.
APPLICANT: Goodearl, Andrew D.J.
TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC AND OTHER
TITLE OF INVENTION: USES
FILE REFERENCE: 07334-325001
CURRENT APPLICATION NUMBER: US/09/803,589
CURRENT FILING DATE: 2001-03-09
PRIOR APPLICATION NUMBER: US 09/128,709
PRIOR FILING DATE: 1998-08-04
PRIOR APPLICATION NUMBER: US 60/054,645
PRIOR FILING DATE: 1997-08-04
PRIOR APPLICATION NUMBER: US 09/130,491
PRIOR FILING DATE: 1998-08-06
PRIOR APPLICATION NUMBER: US 60/054,966
PRIOR FILING DATE: 1997-08-06
PRIOR APPLICATION NUMBER: US 60/058,108
PRIOR FILING DATE: 1997-09-05
PRIOR APPLICATION NUMBER: US 09/388,280
PRIOR FILING DATE: 1999-09-01
PRIOR APPLICATION NUMBER: US 09/388,279
PRIOR FILING DATE: 1999-09-01
NUMBER OF SEQ ID NOS: 14
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10
LENGTH: 518
TYPE: PRT
ORGANISM: Mus musculus
US-09-803-589-10

Query Match 40.9%; Score 45; DB 10; Length 518;
Best Local Similarity 57.1%; Pred. No. 1.5e+02;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 2 NAGSVEWTPKKPGL 15
Db 190 NEPTVETPKYACV 203

RESULT 11
US-10-365-227-16
Sequence 16, Application US/10365227
Publication No. US20030143632A1
GENERAL INFORMATION:
APPLICANT: McCarthy, Sean A.
APPLICANT: Holtzman, Douglas A.
APPLICANT: Goodearl, Andrew D.J.
TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC AND OTHER
TITLE OF INVENTION: USES
FILE REFERENCE: 07334-323001
CURRENT APPLICATION NUMBER: US/10/365,227
CURRENT FILING DATE: 2003-02-12
PRIOR APPLICATION NUMBER: US/09/802,582
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: US 09/128,709
PRIOR FILING DATE: 1998-08-04
PRIOR APPLICATION NUMBER: US 60/054,645
PRIOR FILING DATE: 1997-08-04

; PRIOR APPLICATION NUMBER: US 09/130,491
; PRIOR FILING DATE: 1998-08-06
; PRIOR APPLICATION NUMBER: US 60/054,966
; PRIOR FILING DATE: 1997-08-06
; PRIOR APPLICATION NUMBER: US 60/058,108
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: US 09/388,280
; PRIOR FILING DATE: 1999-09-01
; PRIOR APPLICATION NUMBER: US 09/388,279
; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 551
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-365-227-16

Query Match 40.9%; Score 45; DB 12; Length 551;
Best Local Similarity 57.1%; Pred. No. 1.6e+02;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 2 NAGSVEWTPKKPGL 15
| :||||| |:
Db 223 NEPTVWTPKYAGV 236

RESULT 12
US-10-929-16
; Sequence 16, Application US/10105929
; Publication No. US20020137142A1
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Goodearl, Andrew D.J.
; TITLE OF INVENTION: TANGO-71, TANGO-73, TANGO-74, TANGO-76, AND TANGO-83
; FILE REFERENCE: 09404/041001
; CURRENT APPLICATION NUMBER: US/10/105,929
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/130,491
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-08-07
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/058,108
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/054,961
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 16
; LENGTH: 551
; TYPE: PRT
; ORGANISM: Rattus rattus
US-10-105-929-16

Query Match 40.9%; Score 45; DB 14; Length 551;
Best Local Similarity 57.1%; Pred. No. 1.6e+02;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 2 NAGSVEWTPKKPGL 15
| :||||| |:
Db 223 NEPTVWTPKYAGV 236

RESULT 13
US-09-445-023A-12
; Sequence 12, Application US/09445023A
; Patent No. US20020119167A1
; GENERAL INFORMATION:
; APPLICANT: Hirose, Kunitaka
; APPLICANT: Inoguchi, Eiji
; APPLICANT: Hakozaiki, Michinori
; APPLICANT: Ishioka, Keiko
; APPLICANT: Ishida, Yukako
; APPLICANT: Matsushima, Kouji
; APPLICANT: Kuno, Kouji

; TITLE OF INVENTION Human ADAMTS-1 protein, gene encoding the same, pharmaceutical
; FILE REFERENCE: Q57092
; CURRENT APPLICATION NUMBER: US/09/445,023A
; CURRENT FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: JP 9-160422
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 727
; TYPE: PRT
; ORGANISM: Mus sp.
US-09-445-023A-12

Query Match 40.9%; Score 45; DB 10; Length 727;
Best Local Similarity 57.1%; Pred. No. 2.1e+02;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 2 NAGSVEWTPKKPGL 15
| :||||| |:
Db 399 NEPTVWTPKYAGV 412

RESULT 14
US-10-097-597-12
; Sequence 12, Application US/10097597
; Publication No. US20030022352A1
; GENERAL INFORMATION:
; APPLICANT: Hirose, Kunitaka
; APPLICANT: Inoguchi, Eiji
; APPLICANT: Hakozaiki, Michinori
; APPLICANT: Ishioka, Keiko
; APPLICANT: Ishida, Yukako
; APPLICANT: Matsushima, Kouji
; APPLICANT: Kuno, Kouji
; TITLE OF INVENTION: Human ADAMTS-1 protein, gene encoding the same,
; FILE REFERENCE: Q57092
; CURRENT APPLICATION NUMBER: US/10/097,597
; CURRENT FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: 09/445,023
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: JP 9-160422
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 727
; TYPE: PRT
; ORGANISM: Mus sp.
US-10-097-597-12

Query Match 40.9%; Score 45; DB 15; Length 727;
Best Local Similarity 57.1%; Pred. No. 2.1e+02;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 2 NAGSVEWTPKKPGL 15
| :||||| |:
Db 399 NEPTVWTPKYAGV 412

RESULT 15
US-10-097-580-12
; Sequence 12, Application US/10097580
; Publication No. US20030032168A1
; GENERAL INFORMATION:
; APPLICANT: Hirose, Kunitaka
; APPLICANT: Inoguchi, Eiji
; APPLICANT: Hakozaiki, Michinori
; APPLICANT: Ishioka, Keiko
; APPLICANT: Ishida, Yukako

```

; APPLICANT: Matsushima, Kouji
; APPLICANT: Kuno, Kouji
; TITLE OF INVENTION: Human ADAMTS-1 protein, gene encoding the same, pharmaceutical
; TITLE OF INVENTION: composition and method of immunologically analyzing human ADAMTS
; FILE REFERENCE: Q57092
; CURRENT APPLICATION NUMBER: US/10/097,580
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: 09/445,023
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: JP 9-160422
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 727
; TYPE: PRT
; ORGANISM: Mus sp.
; US-10-097-580-12

```

```

Query Match          40.9%   Score 45;   DB 15;   Length 727;
Best Local Similarity 57.1%   Pred. No. 2.1e+02;
Matches      8;   Conservative    2;   Mismatches    4;   Indels    0;   Gaps    0;

```

```

Qy      2  NAGSVEWTPKKPGL 15
      | :||||| |
Db      399 NEPTVWTPKYGV 412

```

```

Search completed: August 27, 2003, 10:01:01
Job time : 286 secs

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run On: August 27, 2003, 08:17:37 ; Search time 39 Seconds
(without alignments)
24.116 Million cell updates/sec

Title: US-10-042-417A-91

Perfect score: 110

Sequence: 1 NNAGSVEWTPKKPGLRRROT 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	91	82.7	198	1	CDNB_FELCA
2	91	82.7	198	1	CDNB_HUMAN
3	86	78.2	198	1	CDNB_CRIGR
4	74	67.3	197	1	CDNB_MOUSE
5	56	50.9	1271	1	MYPC_CHICK
6	47	42.7	422	1	CPXC_AGRU
7	45	40.9	967	1	ATSL1_RAT
8	45	40.9	968	1	ATSL1_MOUSE
9	44.5	40.5	493	1	PPCK_AERPE
10	44.5	40.5	1110	1	Y256_HUMAN
11	44	40.0	126	1	RS6E_METH
12	44	40.0	129	1	RS6E_ARCFU
13	44	40.0	339	1	CIAL_HUMAN
14	44	40.0	522	1	NFF2_HUMAN
15	43	39.1	136	1	RS6E_METAC
16	43	39.1	136	1	RS6E_METMA
17	43	39.1	514	1	MATK_PLAAE
18	43	39.1	603	1	US26_HCMVA
19	43	39.1	801	1	SUS2_DAUCA
20	43	39.1	805	1	SUS1_SOLTU
21	43	39.1	805	1	SUS2_SOLTU
22	43	39.1	805	1	SUSY_LYCES
23	43	39.1	808	1	SUS1_DAUCA
24	43	39.1	1270	1	MYPC_MOUSE
25	42	38.2	314	1	Y009_HUMAN
26	42	38.2	360	1	CCPH_HSVSA
27	42	38.2	438	1	IAP1_DROME
28	42	38.2	532	1	TCR_ECOLI
29	42	38.2	695	1	TGM2_PAGMA
30	41.5	37.7	276	1	MCT6_MOUSE
31	41	37.3	417	1	WN1A_MOUSE
32	41	37.3	423	1	PG5_MYCPN
33	41	37.3	837	1	GCGR_MOUSE

RESULT 1

CDNB_FELCA

ID CDBN_FELCA STANDARD; PRT; 198 AA.

AC O19001;

DT 15-DEC-1998 (Rel 37, Created)

DT 15-DEC-1998 (Rel 37, Last sequence update)

DE 28-FEB-2003 (Rel 41, Last annotation update)

DE Cyclin-dependent kinase inhibitor 1B (Cyclin-dependent kinase

inhibitor p27) (p27Kip1) (p30 Kip1).

GN CDKN1B OR KIP1.

OS Felis silvestris catus (Cat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.

OX NCBI_TaxID=9685;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Lymph node;

RX MEDLINE=98036042; PubMed=9370275;

RA Okuda M., Minehata K., Setoguchi A., Cho K.-W., Nakamura N.,

RA Nishigaki K., Watari T., Cevario S., O'Brien S.J., Tsujimoto H.,

RA Hasegawa A.;

RT "Cloning and chromosome mapping of the feline genes p21WAF1 and

p27Kip1".

RL Gene 198:141-147(1997).

-!- FUNCTION: Involved in G1 arrest. May mediate TGF beta-induced G1 arrest. Binds to and inhibits complexes formed by cyclin E-CDK2, cyclin A-CDK2, and cyclin D1-CDK4. Interaction with nucleoporin

NUP50 is required for nuclear import and for degradation of

phosphorylated p27Kip1 after nuclear import (By similarity).

-!- SUBUNIT: Interacts with NUP50 (By similarity).

-!- SUBCELLULAR LOCATION: Nuclear (By similarity).

-!- SIMILARITY: THE N-TERMINAL OF CIP1 AND KIP ARE SIMILAR.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; D84649; BRA23167.1; -

InterPro; IPR003175; CDI.

Pfam; PF02234; CIP1; 1.

Cell cycle; Nuclear protein.

FT DOMAIN 153 169

NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).

SQ SEQUENCE 198 AA: 22329 MW; 5A652F81462938D4 CRC64;

Query Match 82.7%; Score 91; DB 1; Length 198;

Best Local Similarity 94.7%; Pred. NO. 3.6e-07;

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 2 NAGSVEWTPKKPGLRRROT 20

||||| |||||||

180 NAGSVEQTPKKPGLRRROT 198

RESULT 2
 CCNDB_HUMAN STANDARD; PRT; 198 AA.
 AC P46527; Q16307; Q9B56;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Cyclin-dependent kinase inhibitor 1B (Cyclin-dependent kinase
 DE inhibitor p27) (p27kip1).
 GN CDKN1B OR KIP1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 28-79 AND 104-152.
 RC MEDLINE=94306518; PubMed=8033212;
 RX Polyak K., Lee M.-H., Erdjument-Bromage H., Koff A., Roberts J.M.,
 RA Tempst P., Massague J.;
 RA "Cloning of p27kip1, a cyclin-dependent kinase inhibitor and a
 RT potential mediator of extracellular antimitogenic signals";
 RL Cell 78:59-66(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95188144; PubMed=7882309;
 RA Pietenpol J.A., Bohlander S.K., Sato Y., Papadopoulos N., Liu B.,
 RA Friedman C., Trask B.J., Roberts J.M., Kinzler K.W., Rowley J.D.;
 RA "Assignment of the human p27kip1 gene to 12p13 and its analysis in
 RT leukemias";
 RL Cancer Res. 55:1206-1210(1995).
 RN [3]
 RP SEQUENCE FROM N.A., AND VARIANTS TRP-15 AND GLY-109.
 RX Rieder M.J., Braun A.C., Montoya M.A., Chung M.-W., Nguyen C.P.,
 RA Nguyen D.A., Livingston R.J., Poel C.L., Robertson P.D.,
 RA Schackwitz W.S., Sherwood J.K., Witrak L.A., Nickerson D.A.;
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A., AND VARIANT GLY-109.
 RX TISSUR=Cervix;
 RA MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Dietzenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fanev J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting J., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RL human and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [5]
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 23-106 OF COMPLEX WITH CDK2
 RP AND CG2A.
 RX MEDLINE=96300318; PubMed=8684460;
 RA Russo A.A., Jeffery P.D., Patten A.K., Massague J., Pavletich N.P.;
 RT "Crystal structure of the p27kip1 cyclin-dependent-kinase inhibitor
 RT bound to the cyclin A-Cdk2 complex.";
 RL Nature 382:325-331(1996).
 CC -!- FUNCTION: Involved in G1 arrest. May mediate TGF beta-induced G1
 CC arrest. Binds to and inhibits complexes formed by cyclin E-CDK2,
 CC cyclin A-CDK2, and cyclin D1-CDK4. Interaction with nucleoporin

CC NUP50 is required for nuclear import and for degradation of
 CC phosphorylated p27kip1 after nuclear import (By similarity).
 CC -!- SUBUNIT: Interacts with NUP50 (By similarity).
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN ALL TISSUES TESTED. HIGHEST
 CC LEVELS IN SKELETAL MUSCLE, LOWEST IN LIVER AND KIDNEY.
 CC -!- DOMAIN: A PEPTIDE SEQUENCE CONTAINING ONLY AA 28-79 RETAINS
 CC SUBSTANTIAL KIP1 CYCLIN A/CDK2 INHIBITORY ACTIVITY.
 CC -!- SIMILARITY: THE N-TERMINAL OF KIP1 AND KIP ARE SIMILAR.
 CC -!- DATABASE: NAME-Atlas Genet. Cytogenet. Oncol. Haematol.;
 CC WWW="http://www.infobiogen.fr/services/Chromancer/Genes/CDKN1BID116.html".
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U10906; AAA20240.1; -;
 DR EMBL; S76988; AAD14244.1; -;
 DR EMBL; S76986; AAD14244.1; JOINED.
 DR EMBL; AF480891; AAL78041.1; -;
 DR EMBL; BC001971; AAO1971.1; -;
 DR PDB; 1J5U; 29-JUL-97.
 DR SWISS-2DPAGE; P46527; HUMAN.
 DR Genew; HGNC:1785; CDKN1B.
 DR GK; P46527; -;
 DR MIM; 600778; -;
 DR GO; GO:0004861; F:cyclin-dependent protein kinase inhibitor a. . .; TAS.
 DR GO; GO:0005072; F:GTPbeta receptor, cytoplasmic mediator acti. . .; TAS.
 DR GO; GO:0007050; P:cell cycle arrest; TAS.
 DR GO; GO:0008285; P:negative regulation of cell proliferation; TAS.
 DR GO; GO:0000079; P:regulation of CDK activity; TAS.
 DR InterPro; IPR003175; CDI.
 DR Pfam; PF02234; CDI; 1.
 KW Cell cycle; Nuclear protein; 3D-structure; Polymorphism.
 FT DOMAIN 153 169 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT VARIANT 15 15 R -> W (in dbSNP:20666828).
 FT VARIANT 109 109 V -> G (in dbSNP:20666827).
 FT CONFLICT 22 22 /FTID=VAR_011872.
 FT E -> D (IN REF. 2).
 SQ SEQUENCE 198 AA; 22073 MW; 1118D58901CDF3FC CRC64;
 Query Match 82.7%; Score 91; DB 1; Length 198;
 Best Local Similarity 94.7%; Pred. No. 3.6e-07;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 2 NAGSVEWTPKKPGLRRRQT 20
 Db 180 NAGSVEQTPKKPGLRRRQT 198
 RESULT 3
 CCNDB_CRICR STANDARD; PRT; 198 AA.
 AC Q60439;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cyclin-dependent kinase inhibitor 1B (Cyclin-dependent kinase
 DE inhibitor p27) (p27kip1) (p30 KIP1).
 GN CDKN1B OR KIP1.
 OS Cricetus griseus (Chinese hamster).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
 OC Cricetus.
 OX NCBI_TaxID=10029;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung;

```

RX MEDLINE-97471701; PubMed-9330642;
RA Parekh H.P., Pillarsetti K., Kunapuli S., Simpkins H.;
RT "Isolation of a hamster cDNA homologous to the mouse and human cyclin
RL kinase inhibitory protein p27Kip1.";
CC
CC -1- FUNCTION: Involved in G1 arrest. May mediate TGF beta-induced G1
CC arrest. Binds to and inhibits complexes formed by cyclin E-CDK2,
CC cyclin A-CDK2, and cyclin D1-CDK4. Interaction with nucleoporin
CC NUP50 is required for nuclear import and for degradation of
CC phosphorylated p27Kip1 after nuclear import.
CC -1- SUBUNIT: Interacts with NUP50 (By similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- SIMILARITY: THE N-TERMINAL OF CIP1 AND KIP ARE SIMILAR.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U49649; AAA2570.1; -
CC InterPro: IPR003175; CDI.
CC Pfam: PF02234; CDI; 1.
CC Cell cycle; Nuclear protein.
CC FT DOMAIN 153 169 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
CC SEQUENCE 198 AA; 22249 MW; 2F3905FAF55EA6E9 CRC64;
CC
CC Query Match 78.2%; Score 86; DB 1; Length 198;
CC Best Local Similarity 89.5%; Pred. No. 2.2e-06;
CC Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
CC
CC QY 2 NAGSVETWPKKPGRLRRQT 20
CC ||||| ||||| ||||| ||
CC Db 180 NAGSVETWPKKPGRLRRQT 198
CC
CC RESULT 4
CCDNB_MOUSE
ID CDNB_MOUSE STANDARD; PRT; 197 AA.
AC P46414;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cyclin-dependent kinase inhibitor 1B (Cyclin-dependent kinase
DE inhibitor p27) (p27Kip1).
GN CDKN1B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-94306519; PubMed-8033213;
RA Toyoshima H., Hunter T.;
RT "p27, a novel inhibitor of G1 cyclin-Cdk protein kinase activity, is
RT related to p21.";
RL Cell 78:67-74(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-94306518; PubMed-8033212;
RA Polyak K., Lee M.-H., Erdjument-Bromage H., Koff A., Roberts J.M.,
RA Tempst P., Massague J.;
RT "Cloning of p27Kip1, a cyclin-dependent kinase inhibitor and a
RT potential mediator of extracellular antimitogenic signals.";
RL Cell 78:59-66(1994).
RN [3]
RP INTERACTION WITH NUP50, AND MUTAGENESIS.
RC STRAIN-BALB/C;
RX MEDLINE-20271857; PubMed-10811608;
RA Mueller D., Thieke K., Buergin A., Dickmanns A., Eilers M.;

```

```

RT "Cyclin E-mediated elimination of p27 requires its interaction with
RT the nuclear pore-associated protein mNAP60.";
RL EMBO J. 19:2168-2180(2000).
CC
CC -1- FUNCTION: Involved in G1 arrest. May mediate TGF beta-induced G1
CC arrest. Binds to and inhibits complexes formed by cyclin E-CDK2,
CC cyclin A-CDK2, and cyclin D1-CDK4. Interaction with nucleoporin
CC NUP50 is required for nuclear import and for degradation of
CC phosphorylated p27Kip1 after nuclear import.
CC -1- SUBUNIT: Interacts with NUP50.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: THE N-TERMINAL OF CIP1 AND KIP ARE SIMILAR.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U10440; AAA21149.1; -
CC PIR: I49064; I49064.
CC MGD: MGI:104565; Cdkn1b.
CC InterPro: IPR003175; CDI.
CC Pfam: PF02234; CDI; 1.
CC Cell cycle; Nuclear protein.
CC KW DOMAIN 153 169 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
CC FT MUTAGEN 90 90 R-SG: LOSS OF INTERACTION WITH NUP50.
CC SQ SEQUENCE 197 AA; 22210 MW; 2D19A6CFE6AE650D CRC64;
CC
CC Query Match 67.3%; Score 74; DB 1; Length 197;
CC Best Local Similarity 82.4%; Pred. No. 0.00017;
CC Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
CC
CC QY 2 NAGSVETWPKKPGRLRR 18
CC ||||| ||||| ||||| ||
CC Db 180 NAGSVETWPKKPGRLRRQ 196
CC
CC RESULT 5
CCMYC_CHICK
ID MYPC_CHICK STANDARD; PRT; 1271 AA.
AC Q90688; Q90907;
DT 15-JUL-1999 (Rel. 38, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Myosin-binding protein C, cardiac-type (Cardiac MyBP-C) (C-protein,
DE cardiac muscle isoform).
GN MYBPC3.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 771-776.
RC TISSUE-Embryonic heart, and Embryonic skeletal muscle;
RX MEDLINE-96129586; PubMed-8576942;
RA Yasuda M., Koshida S., Sato N., Obinata T.;
RT "Complete primary structure of chicken cardiac C-protein (MyBP-C) and
RT its expression in developing striated muscles.";
RL J. Mol. Cell. Cardiol. 27:2275-2286(1995).
RN [2]
RP SEQUENCE FROM N.A.
RA Mohamed A.S., Dignam J.D., Schlender K.K.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF N-TERMINUS, PARTIAL SEQUENCE, AND PHOSPHORYLATION SITES.
RC TISSUE-Heart;
RX MEDLINE-99003479; PubMed-9784245;
RA Mohamed A.S., Dignam J.D., Schlender K.K.;
RT "Cardiac myosin-binding protein C (MyBP-C): identification of protein

```

RT kinase A and protein kinase C phosphorylation sites.";

RL Arch. Biochem. Biophys. 358:313-319(1998).

CC -1- FUNCTION: THICK FILAMENT-ASSOCIATED PROTEIN LOCATED IN THE

CC CROSSBRIDGE REGION OF VERTEBRATE STRIATED MUSCLE. A BANDS. IN VITRO

CC IT BINDS MHC, F-ACTIN AND NATIVE THIN FILAMENTS, AND MODIFIES THE

CC ACTIVITY OF ACTIN-ACTIVED MYOSIN ATPASE. IT MAY MODULATE MUSCLE

CC CONTRACTION OR MAY PLAY A MORE STRUCTURAL ROLE. MAY BE INVOLVED IN

CC THE EARLY PHASE OF MYOFIBRILLOGENESIS.

CC -1- ALTERNATIVE PRODUCTS:

CC Event-Alternative splicing; Named isoforms-2;

CC Name-Type 1;

CC IsoId-Q90688-1; Sequence=Displayed;

CC Name-Type II;

CC IsoId-Q90688-2; Sequence=VSP_002546;

CC -1- TISSUE SPECIFICITY: EXPRESSED SPECIFICALLY IN CARDIAC MUSCLE AMONG

CC ADULT TISSUES, BUT IS ALSO EXPRESSED TRANSIENTLY IN THE SKELETAL

CC MUSCLE AT EARLY DEVELOPMENTAL STAGES. TYPE I ISOFORM IS FOUND IN

CC EMBRYONIC SKELETAL MUSCLE AND TYPE II IS FOUND IN BOTH EMBRYONIC

CC SKELETAL AND CARDIAC MUSCLE.

CC -1- PTM: SUBSTRATE FOR PHOSPHORYLATION BY PKA AND PKC. REVERSIBLE

CC PHOSPHORYLATION APPEARS TO MODULATE CONTRACTION.

CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. MYBP

CC SUBFAMILY.

CC -1- SIMILARITY: Contains 7 immunoglobulin-like C2-type domains.

CC -1- SIMILARITY: Contains 3 fibronectin type III domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL: D43697; BAA07799.1; -

CC EMBL: U38949; AAA92617.1; -

CC HSP; P56276; ITLK.

CC InterPro: IPR003961; FN_III.

CC InterPro: IPR003962; FNIII_subd.

CC InterPro: IPR007110; Ig-like.

CC InterPro: IPR003598; Ig_C2.

CC InterPro: IPR003006; Ig_MHC.

CC Pfam: PF00041; fn3; 3.

CC Pfam: PF00047; Ig; 8.

CC PRINTS: PR00014; FNTPETIII.

CC SMART: SM00060; FN3; 3.

CC SMART: SM00408; IGC2; 1.

CC PROSITE: PS50835; IG-LIKE; 6.

CC Immunoglobulin domain; Actin-binding; Cell adhesion; Muscle protein;

CC Thick filament; Repeat; Phosphorylation; Alternative splicing.

CC INIT_MET 0 0

CC DOMAIN 96 141 PRO-RICH.

CC DOMAIN 136 251 IG-LIKE C2-TYPE 1.

CC DOMAIN 358 450 IG-LIKE C2-TYPE 2.

CC DOMAIN 451 541 IG-LIKE C2-TYPE 3.

CC DOMAIN 542 639 IG-LIKE C2-TYPE 4.

CC DOMAIN 643 762 IG-LIKE C2-TYPE 5.

CC DOMAIN 769 866 FIBRONECTIN TYPE-III 1.

CC DOMAIN 867 967 FIBRONECTIN TYPE-III 2.

CC DOMAIN 968 1056 IG-LIKE C2-TYPE 6.

CC DOMAIN 1063 1177 FIBRONECTIN TYPE-III 3.

CC DOMAIN 1178 1262 IG-LIKE C2-TYPE 7.

CC MOD_RES 264 264 PHOSPHORYLATION (BY PKA AND PKC).

CC MOD_RES 273 273 PHOSPHORYLATION (BY PKA AND PKC).

CC MOD_RES 299 299 PHOSPHORYLATION (BY PKA).

CC MOD_RES 1168 1168 PHOSPHORYLATION (BY PKC).

CC VARSPLOC 264 278 Missing (in isoform Type II).

CC CONFLICT 109 109 MISSING (IN REF. 1).

CC CONFLICT 680 682 IWQ -> SGR (IN REF. 1).

CC CONFLICT 1243 1243 L -> F (IN REF. 1).

CC SEQUENCE 1271 AA; 142156 MW; 4045DF7659C022B4 CRC64;

Query Match 50.9%; Score 56; DB 1; Length 1271;

Best Local Similarity 66.7%; Pred. No. 0.79;

Matches 10; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 GSVETPKPKPGLRRR 18

DB 909 GSAEWTPALPGLTER 923

II IIII IIII I

RESULT 6

CPXC_AGRU STANDARD; PRT; 422 AA.

ID CPXC_AGRU

AC P24466;

DT 01-MAR-1992 (Rel. 21, Created)

DT 01-MAR-1992 (Rel. 21, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DT Cytochrome P450-plnF1, plant-inducible (EC 1.14.-.-).

GN Cytochrome P450-plnF1, plant-inducible (EC 1.14.-.-).

GN CYP103 OR PINF1 OR VIRH1.

OS Agrobacterium tumefaciens.

OG plasmid pTiA6.

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;

OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.

OX NCBI_TaxID=358;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=89213933; PubMed=2708311;

RA Kanamoto R.H., Powell A.T., Akiyoshi D.E., Regier D.A.,

RA Kerstetter R.A., Nester E.W., Hawes M.C., Gordon M.P.;

RT "Nucleotide sequence and analysis of the plant-inducible locus plnF

RL from Agrobacterium tumefaciens.";

RL J. Bacteriol. 171:2505-2512(1989).

CC -1- FUNCTION: NOT ESSENTIAL FOR VIRULENCE, BUT MAY BE INVOLVED IN THE

CC DETOXIFICATION OF PLANT PROTECTIVE AGENTS AT THE SITE OF WOUNDING.

CC -1- INDUCTION: TRANSCRIPTIONALLY ACTIVATED IN THE PRESENCE OF WOUNDED

CC PLANT TISSUE AND BY PLANT PHENOLIC COMPOUNDS, SUCH AS

CC ACETOSYRINGONE.

CC -1- SIMILARITY: Belongs to the cytochrome P450 family.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL: M19352; AAA82502.1; -

CC FIR; A32306; A32306.

CC InterPro: IPR001128; Cytochrome_P450.

CC Pfam: PF00067; P450; 1.

CC PRINTS: PR00385; P450.

CC PROSITE: PS00086; CYTOCHROME_P450; 1.

CC Oxidoreductase; Monooxygenase; Electron transport; Heme; plasmid.

FT METAL 369 369 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).

SEQ SEQUENCE 422 AA; 47519 MW; 6A9FE4AA9B7E2302 CRC64;

Query Match 42.7%; Score 47; DB 1; Length 422;

Best Local Similarity 72.7%; Pred. No. 6.4;

Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 EWTPKPKPGLRR 17

DB 45 EWRPKPFLRR 55

II IIII IIII

RESULT 7

ATSL_RAT STANDARD; PRT; 967 AA.

ID ATSL_RAT

AC Q9WUQ1; Q9ERI1;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE ADAMTS-1 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase

DE with thrombospondin motifs 1) (ADAM-TS 1) (ADAM-TS1).
GN ADAMTS1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Brain;
RA Liu X., Tu Y., Yin T., Johnstone E.M., Stephenson D.T., Clemens J.A.,
RA "Little S.P.;
RT "Induction of a disintegrin and metalloprotease with the
RT thrombospondin type I motif (ADAMTS).";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 18-967 FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Liver;
RX MEDLINE=20304099; PubMed=10847486;
RA Diamantis I., Luechi M., Hoesli M., Reichen J.;
RT "Cloning of the rat ADAMTS-1 gene and its down regulation in
RT endothelial cells in cirrhotic rats.";
RL Liver 20:165-172(2000).
CC -1- FUNCTION: CLEAVES AGGREGAN, A CARTILAGE PROTEOGLYCAN, AND MAY BE
CC INVOLVED IN ITS TURNOVER. HAS ANGIOGENIC INHIBITOR ACTIVITY (BY
CC SIMILARITY). ACTIVE METALLOPROTEASE, WHICH MAY BE ASSOCIATED WITH
CC VARIOUS INFLAMMATORY PROCESSES AS WELL AS DEVELOPMENT OF CANCER
CC CACHEXIA. MAY PLAY A CRITICAL ROLE IN FOLLICULAR RUPTURE (BY
CC SIMILARITY).
CC -1- CATALYTIC ACTIVITY: Cleaves aggrecan at the 1683-Glu-I-Leu-1684
CC site, within the chondroitin sulfate attachment domain.
CC -1- COFACTOR: Binds 1 zinc ion per subunit (by similarity).
CC -1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC matrix (by similarity).
CC -1- INDUCTION: DOWN-REGULATED IN ENDOTHELIAL CELLS DERIVED FROM
CC CIRRHOTIC LIVER.
CC -1- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT
CC FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX.
CC -1- PTM: THE PRECURSOR IS CLEAVED BY A FURIN ENDOPEPTIDASE (BY
CC SIMILARITY).
CC -1- SIMILARITY: Belongs to peptidase family M12B.
CC -1- SIMILARITY: Contains 1 disintegrin-like domain.
CC -1- SIMILARITY: Contains 3 TSP type-1 domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF149118; AAD34012.1; -;
DR EMBL: AF304446; AAG29823.1; -;
DR MEROPS: M12_222; -;
DR InterPro: IPR006586; ADAM_cysteine.
DR InterPro: IPR001762; Disintegrin.
DR InterPro: IPR002870; Pep_M12B_propep.
DR InterPro: IPR001590; Reprolysin.
DR InterPro: IPR000884; TSpl.
DR InterPro: IPR006025; Zn_Mtpetdse.
DR Pfam: PF01562; Pep_M12B_propep; 1.
DR Pfam: PF01421; Reprolysin; 1.
DR Pfam: PF00090; tsp_1; 3.
DR SMART: SM00608; ACR; 1.
DR SMART: SM00209; TSP1; 3.
DR PROSITE: PS00215; ADAM_MEPRO; 1.
DR PROSITE: PS00427; DISINTEGRIN_1; FALSE_NEG.
DR PROSITE: PS00214; DISINTEGRIN_2; FALSE_NEG.
DR PROSITE: PS00142; ZINC_PROTEASE; 1.
DR PROSITE: PS50092; TSpl; 3.
KW Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
KW Repeat; Extracellular matrix; Heparin-binding.
FT SIGNAL 1 54 POTENTIAL.
RN [3]

FT PROPEP 55 252 BY SIMILARITY.
FT CHAIN 253 967 ADAMTS-1.
FT DOMAIN 253 967 METALLOPROTEASE.
FT DOMAIN 476 558 DISINTEGRIN-LIKE.
FT DOMAIN 559 614 TSP TYPE-1 1.
FT DOMAIN 616 724 CYS-RICH.
FT DOMAIN 725 857 SPACER.
FT DOMAIN 854 910 TSP TYPE-1 2.
FT DOMAIN 911 967 TSP TYPE-1 3.
FT DOMAIN 194 198 POLY-ARG.
FT SITE 205 205 CYSTEINE SWITCH (POTENTIAL).
FT METAL 401 401 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 402 402 BY SIMILARITY.
FT METAL 405 405 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 411 411 ZINC (CATALYTIC) (BY SIMILARITY).
FT CARBOHYD 547 547 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 720 720 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 764 764 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 782 782 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 945 945 N-LINKED (GLCNAC...) (POTENTIAL).
FT CONFLICT 21 21 I -> V (IN REF. 2).
FT CONFLICT 26 31 KFRSSQ -> RSRGSL (IN REF. 2).
FT CONFLICT 49 49 V -> A (IN REF. 2).
FT CONFLICT 72 72 R -> P (IN REF. 2).
FT CONFLICT 79 79 L -> TR (IN REF. 2).
FT CONFLICT 249 249 R -> G (IN REF. 2).
FT CONFLICT 262 265 TMLV -> NULLK (IN REF. 2).
FT CONFLICT 607 607 S -> F (IN REF. 2).
FT CONFLICT 936 936 L -> V (IN REF. 2).
FT CONFLICT 962 962 I -> T (IN REF. 2).
SQ SEQUENCE 967 AA: 105705 MW: F93C864F6DCDB4CF CRC64;
Query Match 40.9%; Score 45; DB 1; Length 967;
Best Local Similarity 57.1%; Pred. No. 31;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
QY 2 NAGSVEWTPKKPGL 15
| :||||| |
Db 639 NEPTVEWTPKYAGV 652
RESULT 8
ATSL_MOUSE STANDARD; PRT; 968 AA.
ID ATSL_MOUSE AC P97857; O54768;
DT 30-MAY-2000 (Rel. 39, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE ADAMTS-1 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
DE with thrombospondin motifs 1) (ADAM-TS 1) (ADAM-TS1).
GN ADAMTS1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SvJ;
RX MEDLINE=98110583; PubMed=9441751;
RA Kuno K., Lizasa H., Ohno S., Matsushima K.;
RT "The exon/intron organization and chromosomal mapping of the mouse
RT ADAMTS-1 gene encoding an ADAM family protein with TSP motifs.";
RL Genomics 46:466-471(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97150761; PubMed=8995297;
RA Kuno K., Kanada N., Nakashima E., Fujiki F., Ichimura F.,
RA Matsushima K.;
RT "Molecular cloning of a gene encoding a new type of metalloproteinase-
RT disintegrin family protein with thrombospondin motifs as an
RT inflammation associated gene.";
RL J. Biol. Chem. 272:556-562(1997).
RN [3]

CHARACTERIZATION, AND MUTAGENESIS OF GLU-403.
 MEDLINE=99303657; PubMed=10373500;
 Kuno K., Terashima Y., Matsushima K.;
 "ADAMTS-1 is an active metalloproteinase associated with the
 extracellular matrix.";
 J. Biol. Chem. 274:18821-18826(1999).
 [4]
 FUNCTION.
 MEDLINE=20389568; PubMed=10930576;
 Kuno K., Okada Y., Kawashima H., Nakamura H., Miyasaka M.,
 Ono H., Matsushima K.;
 "ADAMTS-1 cleaves a cartilage proteoglycan, aggrecan.";
 FEBS Lett. 478:241-245(2000).
 [5]
 FUNCTION, AND INDUCTION.
 MEDLINE=20243757; PubMed=10781075;
 Robker R.L., Russell D.L., Espey L.L., Lydon J.P., O'Malley B.W.,
 Richards J.S.;
 "Progesterone-regulated genes in the ovulation process: ADAMTS-1 and
 cathepsin L proteases.";
 Proc. Natl. Acad. Sci. U.S.A. 97:4689-4694(2000).
 CC -1- FUNCTION: CLEAVES AGGECAN, A CARTILAGE PROTEOGLYCAN, AND MAY BE
 INVOLVED IN ITS TURNOVER. HAS ANGIOGENIC INHIBITOR ACTIVITY (BY
 SIMILARITY). ACTIVE METALLOPROTEINASE, WHICH MAY BE ASSOCIATED WITH
 VARIOUS INFLAMMATORY PROCESSES AS WELL AS DEVELOPMENT OF CANCER
 CACHEXIA. MAY PLAY A CRITICAL ROLE IN FOLLICULAR RUPTURE (BY
 SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: Cleaves aggrecan at the 1691-Glu-|-Leu-1692
 site, within the chondroitin sulfate attachment domain.
 CC -1- COFACTOR: binds 1 zinc ion per subunit (By similarity).
 CC -1- SUBCELLULAR LOCATION: SECRETED. ASSOCIATED WITH THE EXTRACELLULAR
 MATRIX.
 CC -1- INDUCTION: INDUCED IN VITRO IN COLON ADENOCARCINOMA CELLS BY
 INTERLEUKIN-1, OR IN VIVO IN KIDNEY AND HEART BY
 LIPOPOLYSACCHARIDE. ALSO INDUCED BY LH STIMULATION IN GRANULOSA
 CELLS OF PREOVULATORY FOLLICLES.
 CC -1- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE-1 DOMAINS ARE IMPORTANT
 FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX.
 CC -1- PTM: THE PRECURSOR IS CLEAVED BY A FURIN ENDOPEPTIDASE.
 CC -1- SIMILARITY: Belongs to peptidase family M12B.
 CC -1- SIMILARITY: Contains 1 disintegrin-like domain.
 CC -1- SIMILARITY: Contains 3 tsp type-1 domains.
 CC -1- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
 FRAMESHIFT IN POSITION 7.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AB001735; BAA24501.1; ALT_INIT.
 DR EMBL; D67076; BAA11088.1; ALT_FRAME.
 DR MEROPS; M12.222; -
 DR MGD; MGI:109249; Adamts1.
 DR InterPro; IPR006586; Adamts1.
 DR InterPro; IPR001762; Disintegrin.
 DR InterPro; IPR002870; Pep_M12B_propep.
 DR InterPro; IPR001590; Reprolysin.
 DR InterPro; IPR000884; TSP1.
 DR InterPro; IPR006025; Zn_MTPeptdse.
 DR Pfam; PF01562; Pep_M12B_propep; 1.
 DR Pfam; PF01421; Reprolysin; 1.
 DR Pfam; PF00090; tsp-1; 3.
 DR SMART; SM00608; ACR; 1.
 DR SMART; SM00209; TSP1; 3.
 DR PROSITE; PS50215; ADAM_MEPRO; 1.
 DR PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
 DR PROSITE; PS50214; DISINTEGRIN_2; FALSE_NEG.
 DR PROSITE; PS50092; TSP1; 3.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.

KW Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
 KW Repeat; Extracellular matrix; Heparin-binding.
 FT SIGNAL 1 48
 FT PROPEP 49 253
 FT CHAIN 254 968
 FT DOMAIN 254 476
 FT DOMAIN 477 559
 FT DOMAIN 560 615
 FT DOMAIN 618 725
 FT DOMAIN 726 850
 FT DOMAIN 855 911
 FT DOMAIN 912 968
 FT DOMAIN 195 199
 FT SITE 206 206
 FT METAL 402 402
 FT ACT_SITE 403 403
 FT METAL 412 412
 FT CARBOHYD 548 548
 FT CARBOHYD 721 721
 FT CARBOHYD 765 765
 FT CARBOHYD 783 783
 FT CARBOHYD 946 946
 FT MUTAGEN 403 403
 FT CONFLICT 335 335
 FT CONFLICT 425 425
 FT SEQUENCE 968 AA; 105841 MW; 42EBDA55499FB6C1 CRC64;
 SQ
 Query Match 40.9%; Score 45; DB 1; Length 968;
 Best Local Similarity 57.1%; Pred. No. 31;
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 QY 2 NAGSVETPKPKCL 15
 Db 640 NEPTVETPKYAGV 653
 | :||||| | :
 RESULT 9
 ID PFCK_AERPE STANDARD; PRT; 493 AA.
 AC QYVG68;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Phosphoenolpyruvate carboxykinase [ATP] (EC 4.1.1.49) (PEP
 carboxykinase) (Phosphoenolpyruvate carboxylase) (PEPCK).
 GN PCKA OR AP50033.
 OS Aeropyrum pernix.
 OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
 OC Desulfurococaceae; Aeropyrum.
 OX NCBI_TaxID=56636;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K1;
 RX MEDLINE=99310339; PubMed=10382966;
 RX Kawanabayasi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
 Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,
 Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
 Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudo Y.,
 Yamazaki J., Kushiida N., Oguchi A., Aoki K.-I., Kubota K.,
 Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
 "Complete genome sequence of an aerobic hyper-thermophilic
 crenarchaeon, Aeropyrum pernix K1";
 DNA Res. 6:83-101(1999).
 RL CC -1- CATALYTIC ACTIVITY: ATP + oxaloacetate = ADP + phosphoenolpyruvate
 + CO(2).
 CC -1- PATHWAY: Rate-limiting gluconeogenic enzyme.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to the phosphoenolpyruvate carboxykinase (ATP)
 family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; AP000038; BAA78942.1; -
DR PIR; D72755; D72755.

DR HAMAP; MF_00453; -; 1.

DR InterPro; IPR001272; PEPCK_ATP.

DR Pfam; PF01293; PEPCK_ATP; 1.

DR PROSITE; PS00532; PEPCK_ATP; FALSE_NEG.

KW Gluconeogenesis; Lyase; Decarboxylase; ATP-binding; Complete proteome.

FT NP_BIND 212 219 ATP (BY SIMILARITY).

SQ SEQUENCE 493 AA; 55518 MW; E9FA70C5E8B8D779 CRC64;

Query Match 40.5%; Score 44.5; DB 1; Length 493;

Best Local Similarity 42.1%; Pred. No. 19;

Matches 8; Conservative 5; Mismatches 5; Indels 1; Gaps 1;

OY 2 NAGSVETPKPKGLRRRQT 20

DB 18 SSADIENPP-PGLLRRES 35

DR Y256_HUMAN STANDARD; PRT; 1110 AA.

AC Q93073;

DT 01-NOV-1997 (Rel. 35, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Hypothetical protein KIAA0256 (Fragment).

GN KIAA0256.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Nagase T., Seki N., Ishikawa K.-I., Ohira M., Kawarabayasi Y.,

RA Ohara O., Tanaka A., Kotani H., Miyajima N., Nomura N.;

RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE OF 476-1110 FROM N.A.

RC TISSUE=Bone marrow;

RX MEDLINE=97191544; PubMed=9039502;

RA Nagase T., Seki N., Ishikawa K.-I., Ohira M., Kawarabayasi Y.,

RA Ohara O., Tanaka A., Kotani H., Miyajima N., Nomura N.;

RT "Prediction of the coding sequences of unidentified human genes. VI.

RT The coding sequences of 80 new genes (KIAA0201-KIAA0280) deduced by

RT analysis of cDNA clones from cell line KG-1 and brain.;"

RL DNA Res. 3:321-329(1996).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; D87445; BAA13386.2; -

DR InterPro; IPR004038; Ribosomal_L7A.

DR Pfam; PF01248; Ribosomal_L7Ae; 1.

KW Hypothetical protein.

FT NON_TER 1

SQ SEQUENCE 1110 AA; 122556 MW; 0499DDA0D2FE5EDF CRC64;

Query Match 40.5%; Score 44.5; DB 1; Length 1110;

Best Local Similarity 40.0%; Pred. No. 43;

Matches 10; Conservative 2; Mismatches 6; Indels 7; Gaps 1;

OY 1

DB 0499DDA0D2FE5EDF CRC64;

DR Y256_HUMAN STANDARD; PRT; 1110 AA.

AC Q93073;

DT 01-NOV-1997 (Rel. 35, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Hypothetical protein.

GN KIAA0256.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Nagase T., Seki N., Ishikawa K.-I., Ohira M., Kawarabayasi Y.,

RA Ohara O., Tanaka A., Kotani H., Miyajima N., Nomura N.;

RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE OF 476-1110 FROM N.A.

RC TISSUE=Bone marrow;

RX MEDLINE=97191544; PubMed=9039502;

RA Nagase T., Seki N., Ishikawa K.-I., Ohira M., Kawarabayasi Y.,

RA Ohara O., Tanaka A., Kotani H., Miyajima N., Nomura N.;

RT "Prediction of the coding sequences of unidentified human genes. VI.

RT The coding sequences of 80 new genes (KIAA0201-KIAA0280) deduced by

RT analysis of cDNA clones from cell line KG-1 and brain.;"

RL DNA Res. 3:321-329(1996).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; D87445; BAA13386.2; -

DR InterPro; IPR004038; Ribosomal_L7A.

DR Pfam; PF01248; Ribosomal_L7Ae; 1.

KW Hypothetical protein.

FT NON_TER 1

SQ SEQUENCE 1110 AA; 122556 MW; 0499DDA0D2FE5EDF CRC64;

Query Match 40.5%; Score 44.5; DB 1; Length 1110;

Best Local Similarity 40.0%; Pred. No. 43;

Matches 10; Conservative 2; Mismatches 6; Indels 7; Gaps 1;

OY 1

DB 0499DDA0D2FE5EDF CRC64;

DR Y256_HUMAN STANDARD; PRT; 1110 AA.

AC Q93073;

DT 01-NOV-1997 (Rel. 35, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Hypothetical protein.

GN KIAA0256.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Nagase T., Seki N., Ishikawa K.-I., Ohira M., Kawarabayasi Y.,

RA Ohara O., Tanaka A., Kotani H., Miyajima N., Nomura N.;

RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE OF 476-1110 FROM N.A.

RC TISSUE=Bone marrow;

RX MEDLINE=97191544; PubMed=9039502;

RA Nagase T., Seki N., Ishikawa K.-I., Ohira M., Kawarabayasi Y.,

RA Ohara O., Tanaka A., Kotani H., Miyajima N., Nomura N.;

RT "Prediction of the coding sequences of unidentified human genes. VI.

RT The coding sequences of 80 new genes (KIAA0201-KIAA0280) deduced by

RT analysis of cDNA clones from cell line KG-1 and brain.;"

RL DNA Res. 3:321-329(1996).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; D87445; BAA13386.2; -

DR InterPro; IPR004038; Ribosomal_L7A.

DR Pfam; PF01248; Ribosomal_L7Ae; 1.

KW Hypothetical protein.

FT NON_TER 1

SQ SEQUENCE 1110 AA; 122556 MW; 0499DDA0D2FE5EDF CRC64;

Query Match 40.5%; Score 44.5; DB 1; Length 1110;

Best Local Similarity 40.0%; Pred. No. 43;

Matches 10; Conservative 2; Mismatches 6; Indels 7; Gaps 1;

OY 1

DB 0499DDA0D2FE5EDF CRC64;

DR Y256_HUMAN STANDARD; PRT; 1110 AA.

AC Q93073;

DT 01-NOV-1997 (Rel. 35, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Hypothetical protein.

GN KIAA0256.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Nagase T., Seki N., Ishikawa K.-I., Ohira M., Kawarabayasi Y.,

RA Ohara O., Tanaka A., Kotani H., Miyajima N., Nomura N.;

RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE OF 476-1110 FROM N.A.

RC TISSUE=Bone marrow;

RX MEDLINE=97191544; PubMed=9039502;

RA Nagase T., Seki N., Ishikawa K.-I., Ohira M., Kawarabayasi Y.,

RA Ohara O., Tanaka A., Kotani H., Miyajima N., Nomura N.;

RT "Prediction of the coding sequences of unidentified human genes. VI.

RT The coding sequences of 80 new genes (KIAA0201-KIAA0280) deduced by

RT analysis of cDNA clones from cell line KG-1 and brain.;"

RL DNA Res. 3:321-329(1996).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; D87445; BAA13386.2; -

DR InterPro; IPR004038; Ribosomal_L7A.

DR Pfam; PF01248; Ribosomal_L7Ae; 1.

KW Hypothetical protein.

FT NON_TER 1

SQ SEQUENCE 1110 AA; 122556 MW; 0499DDA0D2FE5EDF CRC64;

Query Match 40.5%; Score 44.5; DB 1; Length 1110;

Best Local Similarity 40.0%; Pred. No. 43;

Matches 10; Conservative 2; Mismatches 6; Indels 7; Gaps 1;

OY 1

DB 0499DDA0D2FE5EDF CRC64;

DR Y256_HUMAN STANDARD; PRT; 1110 AA.

AC Q93073;

DT 01-NOV-1997 (Rel. 35, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Hypothetical protein.

GN KIAA0256.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Nagase T., Seki N., Ishikawa K.-I., Ohira M., Kawarabayasi Y.,

RA Ohara O., Tanaka A., Kotani H., Miyajima N., Nomura N.;

RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE OF 476-1110 FROM N.A.

RC TISSUE=Bone marrow;

RX MEDLINE=97191544; PubMed=9039502;

RA Nagase T., Seki N., Ishikawa K.-I., Ohira M., Kawarabayasi Y.,

RA Ohara O., Tanaka A., Kotani H., Miyajima N., Nomura N.;

RT "Prediction of the coding sequences of unidentified human genes. VI.

RT The coding sequences of 80 new genes (KIAA0201-KIAA0280) deduced by

RT analysis of cDNA clones from cell line KG-1 and brain.;"

RL DNA Res. 3:321-329(1996).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; D87445; BAA13386.2; -

DR InterPro; IPR004038; Ribosomal_L7A.

DR Pfam; PF01248; Ribosomal_L7Ae; 1.

KW Hypothetical protein.

FT NON_TER 1

SQ SEQUENCE 1110 AA; 122556 MW; 0499DDA0D2FE5EDF CRC64;

Query Match 40.5%; Score 44.5; DB 1; Length 1110;

Best Local Similarity 40.0%; Pred. No. 43;

Matches 10; Conservative 2; Mismatches 6; Indels 7; Gaps 1;

OY 1

DB 0499DDA0D2FE5EDF CRC64;

DR Y256_HUMAN STANDARD; PRT; 1110 AA.

AC Q93073;

DT 01-NOV-1997 (Rel. 35, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Hypothetical protein.


```

RN RP SEQUENCE FROM N.A.
RC STRAIN-VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klench K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Klausner R.D., Kerlavage A.R., Graham D.E., Kyrides N.C.,
RA Felschmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeill L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Uterback T.,
RA Cotton M.D., Spriggs T., Artlich P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus."
RL Nature 390:364-370(1997).
CC -1- SIMILARITY: BELONGS TO THE S6E FAMILY OF RIBOSOMAL PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE001068; AAB90719.1;
CC PIR; G69313; G69313.
CC TIGR; AF0511; -.
CC HAMAP; MF_00512; -.
CC InterPro; IPR001377; Ribosomal_S6E.
CC Pfam; PF01092; Ribosomal_S6e; 1.
CC PROSITE; PS00578; RIBOSOMAL_S6E; 1.
CC Ribosomal protein; Complete proteome.
CC SEQUENCE 129 AA; 14319 MW; ADDEF02714FC48D CRC64;
DR DR
DR PIR; G69313; G69313.
DR TIGR; AF0511; -.
DR HAMAP; MF_00512; -.
DR InterPro; IPR001377; Ribosomal_S6E.
DR Pfam; PF01092; Ribosomal_S6e; 1.
DR PROSITE; PS00578; RIBOSOMAL_S6E; 1.
DR Ribosomal protein; Complete proteome.
KW SEQUENCE 129 AA; 14319 MW; ADDEF02714FC48D CRC64;
SQ SEQUENCE 129 AA; 14319 MW; ADDEF02714FC48D CRC64;

Query Match 40.0%; Score 44; DB 1; Length 129;
Best Local Similarity 47.1%; Pred. No. 5.5;
Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Oy 3 AGSVETPKPKPLRRQ 19
Db 78 SGGVGFPRPKGLRKRK 94

RESULT 13
CIAL_HUMAN
ID CIAL_HUMAN STANDARD; PRT; 339 AA.
AC O76071;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE WD-repeat containing protein Ciao 1.
GN CTAOL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98225157; PubMed=9556563;
RA Johnstone R.W., Wang J., Tommerup N., Vissing H., Roberts T., Shi Y.;
RT "Ciao 1 is a novel WD40 protein that interacts with the tumor
RT suppressor protein Wt1."
RL J. Biol. Chem. 273:10880-10887(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Adams M.D., Loftus B.J., Zhou L., Crosby M., Fuhrmann J., Mason T.M.,
RA Brandon R., Kim U.J., Kerlavage A.R., Venter J.C.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [3]

```

```

RP SEQUENCE FROM N.A.
RC TISSUE=Blood, and Skin;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.C.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP CHARACTERIZATION.
RX MEDLINE=93367400; PubMed=10438340;
RA Johnstone R.W., Tommerup N., Hansen C., Vissing H., Shi Y.;
RT "Structural organization, tissue expression, and chromosomal
RT localization of Ciao 1, a functional modulator of the Wilms' tumor
RT suppressor, Wt1."
RL Immunogenetics 49:900-905(1999).
CC -1- FUNCTION: SEEMS TO SPECIFICALLY MODULATE THE TRANSACTIVATION
CC ACTIVITY OF WT1.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- MISCELLANEOUS: "CIAO" MEANS BRIDGE IN CHINESE.
CC -1- SIMILARITY: Contains 7 WD repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U63810; AAC24948.1;
CC EMBL; AC004020; AAC23493.1;
CC EMBL; BC001395; AAH01395.1;
CC EMBL; BC032812; AAH32812.1;
CC MIM; 604333;
CC GO; GO:0008284; P:positive regulation of cell proliferation; TAS.
CC GO; GO:0006357; P:regulation of transcription from Pol II pro. .; TAS.
CC InterPro; IPR001680; WD40.
CC Pfam; PF00400; WD40; 7.
CC PRINTS; PR00320; GPROTEINRPT.
CC ProDom; PD000018; WD40; 2.
CC SMART; SM00320; WD40; 7.
CC PROSITE; PS00678; WD_REPEATS_1; 1.
CC PROSITE; PSS0082; WD_REPEATS_2; 6.
CC PROSITE; PSS0294; WD_REPEATS_REGION; 1.
CC Repeat; WD repeat; Nuclear protein.
FT REPEAT 14 53 WD 1.
FT REPEAT 59 98 WD 2.
FT REPEAT 103 142 WD 3.
FT REPEAT 148 187 WD 4.
FT REPEAT 192 231 WD 5.
FT REPEAT 250 289 WD 6.
FT REPEAT 301 338 WD 7.
SQ SEQUENCE 339 AA; 37840 MW; 63A8D8257A204FC8 CRC64;
Query Match 40.0%; Score 44; DB 1; Length 339;
Best Local Similarity 70.0%; Pred. No. 15;

```


Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 6 VEWTPKKPGL 15

Db 309 VAWNPKPGL 318

RESULT 14

ID NFF2_HUMAN STANDARD; PRT; 522 AA.
 AC Q9Y5X5; Q96RV1; Q9NR49;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Neuropeptide FF receptor 2 (Neuropeptide G protein-coupled receptor)
 DE (G-protein-coupled receptor HLMAR77).
 GN GPR74 OR NPGPR OR NPFF2.
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1), AND TISSUE SPECIFICITY.
 RC TISSUE=Brain;
 RX MEDLINE=20408933; PubMed=10851242;
 RA Elshourbagy N.A., Ames R.S., Fitzgerald L.R., Foley J.J.,
 RA Chambers J.K., Szekeres P.G., Evans N.A., Schmidt D.B., Buckley P.T.,
 RA Dytko G.M., Murodock P.R., Milligan G., Groarke D.A., Tan K.B.,
 RA Shabon U., Nuthulaganti P., Wang D.Y., Wilson S., Bergsma D.J.,
 RA Sarau H.M.;
 RT "Receptor for the pain modulatory neuropeptides FF and AF is an orphan
 RT G protein-coupled receptor.";
 RL J. Biol. Chem. 275:25965-25971(2000).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 2), AND CHARACTERIZATION.
 RC TISSUE=Brain;
 RX MEDLINE=20564301; PubMed=11024015;
 RA Bonini J.A., Jones K.A., Adham N., Forray C., Artymyshyn R.,
 RA Durkin M.M., Smith K.E., Tamm J.A., Boteju L.W., Lakhani P.P.,
 RA Raddatz R., Yao W.-J., Ogozalek K.L., Boyle N., Kouranova E.V.,
 RA Qian Y., Vaysse P.J., Wetzel J.M., Branchek T.A., Gerald C.,
 RA Borowsky B.;
 RT "Identification and characterization of two G protein-coupled
 RT receptors for neuropeptide FF.";
 RL J. Biol. Chem. 275:39324-39331(2000).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM 3).
 RX MEDLINE=20291943; PubMed=10837915;
 RA Parker R.M.C., Copeland N.G., Eyre H.J., Liu M., Gilbert D.J.,
 RA Crawford J., Couzens M., Sutherland G.R., Jenkins N.A., Herzog H.;
 RT "Molecular cloning and characterisation of GPR74 a novel G-protein
 RT coupled receptor closely related to the Y-receptor family.";
 RL Brain Res. Mol. Brain Res. 77:199-208(2000).
 RN [5]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RA Liu Q., Guan X.-M., McDonald T.P., Jiang Q., Zeng Z., Marlene J.,
 RA Williams D.L. Jr., Hong Y., Figueroa D., Clements M.K., Mallee J.,
 RA Wang R., Evans J., Gould R., Austin C.P.;
 RT "Identification and characterization of two cognate receptors for
 RT mammalian FMRFamide-like neuropeptides.";
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A. (ISOFORM 4).
 RA Laemmle B.S., Schindler M., Beilmann M., Hamilton B.S., Doods H.N.,
 RA Wieland H.A.;

RT "Cloning and characterization of the NPGP receptor and identification
 RT of a novel short mRNA isoform in human hypothalamus.";
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: RECEPTOR FOR NPAF (A-18-F-AMIDE) AND NPFF (F-8-F-AMIDE).
 CC NEUROPEPTIDES, ALSO KNOWN AS MORPHINE-MODULATING PEPTIDES. CAN
 CC ALSO BE ACTIVATED BY A VARIETY OF NATURALLY OCCURRING OR SYNTHETIC
 CC FMRF-AMIDE LIKE LIGANDS. THIS RECEPTOR MEDIATES ITS ACTION BY
 CC ASSOCIATION WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-
 CC CALCIUM SECOND MESSENGER SYSTEM.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing: Named isoforms=4;
 CC Comment-Experimental confirmation may be lacking for some
 CC isoforms;
 CC Name=1; Synonyms=Long form;
 CC IsoId=Q9Y5X5-1; Sequence=Displayed;
 CC Name=2; Synonyms=short form;
 CC IsoId=Q9Y5X5-2; Sequence=VSP_001907;
 CC Name=3;
 CC IsoId=Q9Y5X5-3; Sequence=VSP_001908, VSP_001909;
 CC Name=4;
 CC IsoId=Q9Y5X5-4; Sequence=VSP_001910, VSP_001911;
 CC -!- TISSUE SPECIFICITY: Isoform 1 is abundant in placenta. Relatively
 CC highly expressed in thymus, testis, and small intestine. Expressed
 CC at low levels in several tissues including spleen, prostate,
 CC brain, heart, ovary, colon, kidney, lung, liver and pancreas and
 CC not expressed in skeletal muscle and leukocytes. Highest but
 CC relatively low level of isoform 2 in placenta and very low level
 CC in numerous tissues including adipose tissue and many brain
 CC regions. Isoform 3 is expressed in brain and heart and, at lower
 CC levels, in kidney, liver, lung and pancreas.
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC HIGHEST TO OREXIN, NPY AND CHOLECYSTOKININ RECEPTORS.
 CC -!- CAUTION: Ref.1 sequence differs from that shown due to a
 CC frameshift in position 503.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF119815; AM22047.1; -;
 CC EMBL; AF257210; AAF87078.1; -;
 CC EMBL; AF268899; AAG41398.1; -;
 CC EMBL; AF236083; AAK58513.1; ALT_FRAME.
 CC EMBL; AF330053; AAK94197.1; -;
 CC EMBL; AJ311393; CAC85427.1; -;
 CC Genew; HGNC:4525; GPR74.
 CC MIM; 607449; -;
 CC GO; GO:0005887; C:integral to plasma membrane; TAS.
 CC GO; GO:0008151; P:cell growth and/or maintenance; TAS.
 CC GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; TAS.
 CC GO; GO:0009582; P:perception of abiotic stimulus; TAS.
 CC InterPro; IPR000276; GPCR_Rhodpsn.
 CC InterPro; IPR005395; NPFF_receptor.
 CC InterPro; IPR005397; NPFF_receptor2.
 CC Pfam; PF00001; 7tm_1; 1.
 CC PRINTS; PR00237; GPCRHHODOPS.
 CC PRINTS; PR01570; NPFFRECEPTOR.
 CC PRINTS; PR01572; NPFFRECEPTOR.
 CC PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
 CC PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
 CC G-protein coupled receptor; Transmembrane; Glycoprotein;
 CC Alternative splicing.
 CC DOMAIN 1 147 EXTRACELLULAR (POTENTIAL).
 CC TRANSMEM 148 168 1 (POTENTIAL).
 CC DOMAIN 169 184 CYTOPLASMIC (POTENTIAL).
 CC TRANSMEM 185 205 2 (POTENTIAL).
 CC DOMAIN 206 221 EXTRACELLULAR (POTENTIAL).
 CC TRANSMEM 222 242 3 (POTENTIAL).

```

FT DOMAIN 243 262 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 263 283 4 (POTENTIAL).
FT DOMAIN 284 319 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 320 340 5 (POTENTIAL).
FT DOMAIN 341 377 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 378 398 6 (POTENTIAL).
FT DOMAIN 399 413 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 414 434 7 (POTENTIAL).
FT DOMAIN 435 522 CYTOPLASMIC (POTENTIAL).
FT DRSLFID 220 308 BY SIMILARITY.
FT CARBOHYD 110 110 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 122 122 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 133 133 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 300 300 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 1 102 Missing (in isoform 2).
FT VARSPLIC 1 99 /FTid=vsp_001907.
FT VARSPLIC 100 100 Missing (in isoform 3).
FT VARSPLIC 101 132 R -> M (in isoform 3).
FT VARSPLIC 101 132 FIMNEKWDNTSSNHPINWVNTKHHLYSDI -> MAIWK
FT HDVQDQWIGPNCISFSLVSCNCR (in isoform
4).
FT VARSPLIC 133 522 /FTid=vsp_001910.
FT VARSPLIC 466 466 Missing (in isoform 4).
FT CONFLICT 466 466 /FTid=vsp_001911.
FT SEQUENCE 522 AA; 60270 MW; 40CB9FCD42F77041 CRC64;
A -> T (IN REF. 1 AND 4).

Query Match 40.0%; Score 44; DB 1; Length 522;
Best Local Similarity 50.0%; Pred. No. 24;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 3 AGSVETPKKPGGLRRR 18
DB 45 SGSLEWSRQSGDRRR 60

RESULT 15
RS6E_METAC
ID RS6E_METAC STANDARD; PRT; 136 AA.
AC Q8TQL4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 30S ribosomal protein S6e.
GN RPS6E OR MA1526.
OS Methanosarcina acetivorans.
OC Archaea; Euryarchaeota; Euryarchaeota orders incertae sedis;
OC Methanosarcinales; Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2214;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C2A / ATCC 35395 / DSM 2834;
RX MEDLINE=21929760; PubMed=11932238;
RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,
RA FitzHugh W., Calvo S., Engels R., Smirnov S., Atnoor D., Brown A.,
RA Linton N., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,
RA Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,
RA Hedderich R., Ingram-Smith C., Kuetner H.C., Krzycki J.A.,
RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
RA Perry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
RA Metcalf W.W., Birren B.;
RT "The genome of Methanosarcina acetivorans reveals extensive metabolic
and physiological diversity.";
RL Genome Res. 12:532-542(2002).
CC -1- SIMILARITY: BELONGS TO THE S6E FAMILY OF RIBOSOMAL PROTEINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE010822; AAM04940.1; -
CC DR HAMAP; MF_00512; -; 1
CC DR InterPro; IPR001377; Ribosomal_S6E.
CC DR Pfam; PF01092; Ribosomal_S6e; 1.
CC DR PROSITE; PS00578; RIBOSOMAL_S6E; FALSE_NEG.
CC KW Ribosomal protein; Complete proteome.
CC SQ SEQUENCE 136 AA; 14455 MW; 984E265E6276770A CRC64;

Query Match 39.1%; Score 43; DB 1; Length 136;
Best Local Similarity 52.9%; Pred. NO. 8.3;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 AGSVETPKKPGGLRRRQ 19
DB 78 ATGVGVKLPQGGRRR 94
| : ||| || |||
| : ||| || |||

Search completed: August 27, 2003, 09:41:58
Job time : 41 secs

```

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	104	94.5	19	23	AA022505	Human p27 terminus h	
2	91	82.7	19	21	AA083097	Human p27 peptide	
3	91	82.7	19	23	AA022502	Carboxy terminus o	
4	91	82.7	25	22	AA848274	Carboxy terminal e	
5	91	82.7	198	16	AA079133	Human kipl, p27 -ki	
6	91	82.7	198	17	AA922709	Human p27 Kipl. H	
7	91	82.7	198	17	AA927917	27 kDa protein inh	
8	91	82.7	198	19	AA046888	Amino acid sequenc	
9	91	82.7	198	20	AA000768	CKI/KIP protein p2	

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

CC screening compounds for the treatment of proliferative or differentiative
 CC disorders, particularly cancer. These compounds include small molecules,
 CC or compounds or derivatives or analogues of the new ubiquitin ligases.
 CC The compounds are useful for treating diseases such as cancer (e.g.
 CC breast cancer, prostate cancer or ovarian cancer, lymphoma, small cell
 CC lung carcinoma or parathyroid adenomas), major opportunistic infections,
 CC immune disorders, certain cardiovascular diseases or inflammatory
 CC disorders. This sequence represents a peptide of an F-box protein (FBP)
 CC relating to the invention.

XX SQ Sequence 19 AA;

Query Match 82.7%; Score 91; DB 23; Length 19;
 Best Local Similarity 94.7%; Pred. No. 7.8e-07;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 NAGSVEVTPKKGLRRRQT 20
 Db 1 NAGSVEVTPKKGLRRRQT 19

RESULT 4
 AAB48274
 ID AAB48274 standard; peptide; 25 AA.

XX AC AAB48274;

XX DT 02-APR-2001 (first entry)

XX DE Carboxy terminal end of p27 peptide.

XX KW S-phase kinase associated protein; SKP1; SKP2; SKP2-like protein;
 KW CUL-1; cullin; CDC53; p27; cyclin E; Max; Mad; c-Myc; MDM2; p53; Bax;
 KW Bad; Bcl-2; tumour; cytostatic.

XX OS Homo sapiens.

XX XN WO200075184-A1.

XX PD 14-DEC-2000.

XX PF 05-JUN-2000; 2000WO-US15449.

XX PR 04-JUN-1999; 99US-0137494.

XX PA (UYXA) UNIV YALE.

XX PI Zhang H, Tsvetkov LM, Kondo T;

XX DR WPI; 2001-061703/07.

XX PT Modulating polypeptide levels in a cell, diagnosing and treating tumor,
 PT involves altering levels of proteins such as S-phase kinase associated
 PT proteins 1, 2 and cullin/CDC53 proteins -

XX PS Claim 13; Page 49; 162pp; English.

XX CC The invention relates to methods of altering the polypeptide levels in a
 CC cell, using proteins selected from S-phase kinase associated proteins 1
 CC and 2 (SKP1, SKP2). SKP2-like protein and CUL-1 (a member of the cullin/
 CC CDC53 family of proteins). The method is useful for altering the level of
 CC p27, cyclin E, Max, Mad, c-Myc, MDM2, p53, Bax, Bad or Bcl-2 polypeptide
 CC in a cell. SKP2 and SKP2-like protein levels are useful for detecting
 CC tumours, and in monitoring tumor treatment in a mammal. Agents that
 CC modulate interactions between SKP and target proteins are useful for
 CC treating tumours. The present sequence represents a p27 carboxy terminal
 CC peptide.

XX SQ Sequence 25 AA;

Query Match 82.7%; Score 91; DB 22; Length 25;

Best Local Similarity 94.7%; Pred. No. 1e-06;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 NAGSVEVTPKKGLRRRQT 20
 Db 7 NAGSVEVTPKKGLRRRQT 25

RESULT 5
 AAR79133

ID AAR79133 standard; protein; 198 AA.

XX AC AAR79133;

XX DT 25-MAR-2003 (updated)

XX DT 27-FEB-1996 (first entry)

XX DE Human kipl, p27-kipl or p27, cyclin E-Cdk2 activation inhibitor.

XX KW Human kipl; p27-kipl; p27; cyclin E-Cdk2 complex; cancer;
 KW activation inhibitor; hyperplasia; cyclin dependent kinase;
 KW diagnosis; hyperproliferative disorder; ulcer; partial protein.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers
 FT Peptide 31..43

FT /note= "obtd. from purified kipl, and
 FT used to design degenerate oligo-
 FT nucleotide PCR primer"

FT Peptide 74..79

FT /note= "obtd. from purified kipl, and.
 FT used to design degenerate oligo-
 FT nucleotide PCR primer"

FT Peptide 83..96

FT /note= "obtd. from purified kipl"

FT Peptide 114..122

FT /note= "obtd. from purified kipl"

FT Peptide 135..147

FT /note= "obtd. from purified kipl"

XX XN WO9518824-A1.

XX PD 13-JUL-1995.

XX PF 09-JAN-1995; 95WO-US00247.

XX PR 15-JUL-1994; 94US-0275983.

XX PR 07-JAN-1994; 94US-0179045.

XX PA (SLOK) SLOAN KETTERING INST CANCER RES.
 XX (HUTC-) HUTCHINSON CANCER RES CENT FRED.

XX PI Koff A, Massague J, Polyak K, Roberts JM;

XX DR WPI; 1995-255037/33.

XX PT Novel protein, p27 inhibits activation of a cyclin E-Cdk2 complex -
 XX useful for treatment of hyper-proliferative disorders, e.g. cancer,
 XX hyperplasia or ulcers.

XX PS Disclosure; Fig 9A; 121pp; English.

XX CC AAR79131-R79133 are respective mink, murine and human kipl (p27-kipl
 CC or p27) partial proteins. p27 inhibits the activation of the cyclin
 CC E-Cdk2 (cyclin dependent kinase) complex. Agents which inhibit or
 CC enhance the ability of p27 to inhibit the activation of cyclin
 CC E-Cdk2, can be used to treat hyperproliferative disorders, e.g.
 CC cancer, hyperplasia or ulcer. Diagnosis of hyperproliferative
 CC disorders, esp. human cancer, can be achieved by detecting a p27
 CC mutation in the cells of the patient. The disorder can be treated
 CC using a pharmaceutical compsn. comprising a recombinant virus
 CC contg. a nucleic acid mol. encoding p27.

XX CC (Updated on 25-MAR-2003 to correct PA field.)

```

SQ Sequence 198 AA;
Query Match 82.7%; Score 91; DB 16; Length 198;
Best Local Similarity 94.7%; Pred. No. 8.5e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 NAGSVETPKPKGLRRRQT 20
    ||||| |||||
Db 180 NAGSVETPKPKGLRRRQT 198

RESULT 6
AAW92709
ID AAR92709 standard; Protein; 198 AA.
XX AC AAR92709;
XX DT 16-JUL-1996 (first entry)
XX DE Human p27 Kip1.
XX KW p27 protein; Kip1; cyclin E; Cdk2; cell proliferation; ulcer;
XX KW cancer; hyperplasia; diagnosis; therapy.
XX OS Homo sapiens.
XX PN WO9602140-A1.
XX PD 01-FEB-1996.
XX PF 07-JUN-1995; 95WO-US07361.
XX PR 15-JUL-1994; 94US-0275983.
XX PA (HUTC-) HUTCHINSON CANCER RES CENT FRED.
XX PA (SLOK ) SLOAN RETTERING INST CANCER RES.
XX PI Koff A, Massague J, Polyak K, Roberts JM;
XX WPI; 1996-105553/11.
XX DR N-PSDB; AAT16336.
XX CC p27, an inhibitor of cyclin E-Cdk2 complex activation - and agents
XX PT which enhance and inhibit its activity, useful for treating
XX PT hyper-proliferative and hypo-proliferative disorders
XX PS Disclosure; Fig 15a-b; 129pp; English.
XX CC A cDNA clone (AAT16336) derived from human kidney codes for a 27
XX CC kDa protein, p27 Kip1 (AAR92709), that is capable of binding to
XX CC and inhibiting the activation of a cyclin E-Cdk2 complex. Human
XX CC Kip1 shows a high degree of homology to mink (AAR92707) and mouse
XX CC (AAR92708) Kip1 proteins, and the N-terminal half of the protein shows
XX CC significant homology to Cip1/WAF1. Kip1 shows cdk inhibitory
XX CC activity and prevents cdk2 activation. Overexpression inhibits
XX CC cell entry to the S phase. Kip1 can be produced by expression of
XX CC the cDNA clone in cultured cells. It can be used in vitro
XX CC assays to screen agents that affect p27 activity, and in methods
XX CC for the diagnosis and treatment of hyperproliferative disorders,
XX CC e.g. ulcer, and hyperproliferative disorders, e.g. cancer and
XX CC hyperplasia.

SQ Sequence 198 AA;
Query Match 82.7%; Score 91; DB 17; Length 198;
Best Local Similarity 94.7%; Pred. No. 8.5e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 NAGSVETPKPKGLRRRQT 20
    ||||| |||||
Db 180 NAGSVETPKPKGLRRRQT 198

RESULT 7
AAW29717
ID AAW29717 standard; Protein; 198 AA.
XX AC AAW29717;
XX DT 27-OCT-1998 (first entry)
XX DE 27 kDa protein inhibiting activation of cyclin E-Cdk2 complex.
XX KW 27 kDa protein; p27; kip1; inhibit; activation; cancer; breast carcinoma;
XX KW cyclin E-cyclin-dependent kinase2 complex; hyperproliferative disease;
XX KW recurrence; treatment; human.
XX OS Homo sapiens.
XX PN WO9834121-A2.
XX PD 06-AUG-1998.
XX PF 03-FEB-1998; 98WO-US01893.
XX PR 03-FEB-1997; 97US-0794002.
XX PA (HUTC-) HUTCHINSON CANCER RES CENT FRED.
XX PI Porter PL, Roberts JM;
XX WPI; 1998-437612/37.
XX DR N-PSDB; AAV47517.
XX CC Assays for protein p27 inhibiting activation of cyclin E-Cdk2
XX PT complex - useful for, e.g. diagnosis and prognosis of cancer,
XX PT especially breast carcinoma
XX PS Claim 18; Fig 15B; 105pp; English.
XX CC The present sequence represents a 27 kDa protein (p27 or Kip1) which
XX CC inhibits the activation of a cyclin E-cyclin-dependent kinase2 (Cdk2)
XX CC complex. A reduced relative level of kip1 is indicative of a
XX CC hyperproliferative disease (particularly cancer, especially breast
XX CC carcinoma) and also is prognostic for increased risk of death and/or
XX CC recurrence of cancer (and may be used to determine suitable treatments).
XX CC Agents that affect the activity of kip1 can be used to treat
XX CC hyperproliferative conditions, e.g. to stimulate tissue or organ repair
XX CC or to establish cell cultures.

SQ Sequence 198 AA;
Query Match 82.7%; Score 91; DB 19; Length 198;
Best Local Similarity 94.7%; Pred. No. 8.5e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 NAGSVETPKPKGLRRRQT 20
    ||||| |||||
Db 180 NAGSVETPKPKGLRRRQT 198

RESULT 8
AAW46888
ID AAW46888 standard; Protein; 198 AA.
XX AC AAW46888;
XX DT 25-MAR-2003 (updated)
XX DT 15-JUN-1998 (first entry)
XX DE Amino acid sequence of the p27Kip1 protein.
XX KW E7 oncoprotein; proliferative state; HPV; kinase activity;
XX KW cyclin/cyclin-dependent kinase; p21Cip1; interaction; inactivation;
XX KW cyclin/cyclin-dependent kinase inhibitor.
XX
```



```

PR 09-APR-1999; 99US-0128515.
XX (CELL-) CELL GENESYS INC.
PA (MITO-) MITOTIX INC.
XX
XX Patel S, McArthur J, Gyuris J, Mendez MJ, Finer M;
PI WPI; 2000-587315/55.
XX N-PSDB; AAA90920.
DR
XX Protein and nucleic acid compositions for preventing and treating
PT neoplasias (particularly cancer), comprises a novel chimeric cyclin
PT dependent kinase inhibitor and adenovirus E4 protein -
XX
XX Example 1; Page 122; 126pp; English.
XX
XX This sequence represents the human p27 protein.
CC The invention relates to a protein composition comprising a novel
CC purified chimeric cyclin dependent kinase inhibitor (CDKi) and a
CC purified adenovirus E4 protein. The compositions comprising the protein,
CC or the DNA encoding it are useful for treating neoplasias in animals. The
CC compositions also find use in assays to eliminate a specific
CC sub-population of cultured cells, to determine the susceptibility of
CC neoplastic cells to treatment with the compositions and also in assays to
CC synchronise cell growth in cultured cells.
XX
XX Sequence 198 AA;
SQ
    Query Match      82.7%; Score 91; DB 21; Length 198;
    Best Local Similarity 94.7%; Pred. No. 8.5e-06;
    Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 NAGSVETPKKPGRLRRQT 20
   ||||| ||||| |||||
Db 180 NAGSVETPKKPGRLRRQT 198

RESULT 11
AAY96052
ID AAY96052 standard; Protein; 198 AA.
XX
XX AAY96052;
XX
XX 05-DEC-2000 (first entry)
XX Human cyclin dependent kinase inhibitor p27.
XX
XX Cyclin dependent kinase inhibitor; CDKi; CIP; KIP; human; p27;
XX smooth muscle cell; restenosis; vasotropic; antiproliferative;
XX gene therapy.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Domain 25..93
XX Domain /note= "CDK inhibitory domain"
XX Domain 144..194
XX Peptide /note= "QT domain"
XX Peptide 152..166
XX Modified-site /note= "nuclear localisation signal"
XX Modified-site 10..13
XX Modified-site /note= "O-phosphorylated; weak CDK phosphorylation
XX site"
XX Modified-site 178..181
XX Modified-site /note= "O-phosphorylated; phosphorylation site for
XX proline-directed kinases"
XX Modified-site 187..190
XX Modified-site /note= "O-phosphorylated; CDK phosphorylation
XX consensus site"
XX
XX WO200052159-A1.
XX
XX 08-SEP-2000.

```

```

XX 28-FEB-2000; 2000WO-US04971.
XX
XX 01-MAR-1999; 99US-0122974.
XX 05-NOV-1999; 99US-0163682.
XX 09-DEC-1999; 99US-0457568.
XX
XX (CELL-) CELL GENESYS INC.
XX (MITO-) MITOTIX INC.
XX
XX McArthur J, Gyuris J, Finer M;
XX WPI; 2000-594183/56.
XX N-PSDB; AAA50499.
XX
XX Novel recombinant lentivirus for inhibiting proliferation of smooth
XX muscle cells in e.g. restenosis, is replication deficient and comprises
XX a transgene encoding a cyclin dependent kinase inhibitor -
XX
XX Example 1; Page 119-120; 126pp; English.
XX
XX The present sequence is that of human p27, a cyclin dependent
XX kinase inhibitor (CDKi) that inhibits smooth muscle cell
XX proliferation. A claimed method for inhibiting smooth muscle cell
XX hyperproliferation involves transducing smooth muscle cells with a
XX replication-deficient recombinant adenovirus that lacks a functional
XX E1 region and a functional E4 region, and comprises a transgene
XX encoding a CDKi. The CDKi is selected from an INK4 family protein
XX such as human p16, a CIP/KIP family protein such as p27, active
XX fragments of these, or fusion proteins comprising (active fragments
XX of) an INK4 family protein and a CIP/KIP family protein (see AAY96046
XX and AAY96049). The method is used to inhibit mammalian smooth muscle
XX cell hyperproliferation induced by injury caused by angioplasty,
XX stent placement or vein engraftment. It is useful for treating
XX vascular pathologies, e.g. restenosis. Also claimed are recombinant
XX lentiviruses encoding CDKis.
XX
XX Sequence 198 AA;
XX
XX Query Match      82.7%; Score 91; DB 21; Length 198;
XX Best Local Similarity 94.7%; Pred. No. 8.5e-06;
XX Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 NAGSVETPKKPGRLRRQT 20
   ||||| ||||| |||||
Db 180 NAGSVETPKKPGRLRRQT 198

RESULT 12
AAY96066
ID AAY96066 standard; Protein; 198 AA.
XX
XX AAY96066;
XX
XX 05-DEC-2000 (first entry)
XX Human cyclin dependent kinase inhibitor p27.
XX
XX Cyclin dependent kinase inhibitor; CDKi; CIP; KIP; human; p27;
XX angiogenesis; inhibitor; neoplasia; rheumatoid arthritis;
XX endometriosis; psoriasis; vascular retinopathy; cytostatic;
XX antiarthritic; antirheumatic; gynaecological; antipsoriatic;
XX antiproliferative; gene therapy.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Domain 25..93
XX Domain /note= "CDK inhibitory domain"
XX Domain 144..194
XX Peptide /note= "QT domain"
XX Peptide 152..166
XX /note= "nuclear localisation signal"

```


FT Modified-site 10..13
 FT /note= "O-phosphorylated; weak CDK phosphorylation
 FT site"
 FT Modified-site 178..181
 FT /note= "O-phosphorylated; phosphorylation site for
 FT proline-directed kinases"
 FT Modified-site 187..190
 FT /note= "O-phosphorylated; CDK phosphorylation
 FT consensus site"
 PN WO200052158-A1.
 XX
 PD 08-SEP-2000.
 XX
 XX 28-FEB-2000; 2000WO-US04970.
 XX
 PR 01-MAR-1999; 99US-0122974.
 PR 05-NOV-1999; 99US-0163682.
 PR 09-DEC-1999; 99US-0457646.
 XX
 XX (CELL-) CELL GENESYS INC.
 PA (MITO-) MITOTIX INC.
 PI Patel S, Mearthur J, Gyuris J;
 XX
 XX WPI; 2000-565501/52.
 DR N-PSDB; AAA50519.
 XX
 XX Inhibiting angiogenesis and treating angiogenesis-associated
 PT conditions, e.g. neoplasia, psoriasis by transducing an endothelial
 PT cell with a recombinant virus having a transgene encoding a cyclin
 PT dependent kinase inhibitor
 XX
 PS Example 1; Page 127-128; 138pp; English.
 XX
 CC The present sequence is that of human p27, a cyclin dependent
 CC kinase inhibitor (CKI) that inhibits angiogenesis. A claimed
 CC method for inhibiting angiogenesis involves transducing an
 CC epithelial cell with a transgene encoding (internalizable,
 CC secretable) CDK1. The delivery system for the transgene may be a
 CC liposome or a recombinant virus. The CDK1 is preferably a protein
 CC of the CIP/KIP family such as p27, a protein of the INK4 family
 CC such as p16, active fragments of these proteins (e.g. amino acids
 CC 25-93 or 12-178 of human p27), or a fusion of 2 CDK1 proteins such
 CC as p27 and p16 (see AAY96068-80). The method is useful in treating
 CC conditions associated with angiogenesis, e.g. neoplasia, rheumatoid
 CC arthritis, endometriosis, psoriasis and vascular retinopathy
 CC (claimed). Alternatively, the transgene is delivered to an
 CC auxiliary cell, and is expressed by that cell such that the CDK1 is
 CC released into the blood and contacts the target epithelial cell.
 XX
 SQ Sequence 198 AA;
 Query Match 82.7%; Score 91; DB 21; Length 198;
 Best Local Similarity 94.7%; Pred. No. 8.5e-06;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 2 NAGSVEWTPKKPGLRRQT 20
 ID AAY70307
 Db 180 NAGSVEQTPKKPGLRRQT 198
 RESULT 13
 AAY70307
 ID AAY70307 standard; Protein; 198 AA.
 XX
 AC AAY70307;
 XX
 XX 06-JUN-2000 (first entry)
 DT Human mutant cyclin-dependent kinase inhibitor (CKI), p27 S10A protein.
 XX
 DE Cyclin-dependent kinase inhibitor; CKI; p27; human; mutant; G1 phase;
 KW

KW KIS; serine/threonine kinase; cell proliferation; modulator; treatment;
 KW cell proliferative disease; vascular disorder; gene therapy; restenosis;
 KW atherosclerosis.
 XX Homo sapiens.
 OS
 XX Key Location/Qualifiers
 FH Misc-difference 10
 FT /note= "Wild type Ser substituted with Ala"
 FT
 XX WO200011165-A1.
 PN
 XX 02-MAR-2000.
 PD
 XX 20-AUG-1999; 99WO-US18903.
 XX
 PR 21-AUG-1998; 98US-0097710.
 XX
 PA (NABE/) NABEL G J.
 PA (NABE/) NABEL E G.
 XX
 XX Nabel GJ, Nabel EG;
 PI
 XX WPI; 2000-237648/20.
 DR N-PSDB; AAZ51357.
 DR
 XX Novel serine/threonine kinase hKIS polynucleotides and polypeptides
 PT used for inhibiting the cyclin kinase inhibitor p27, and so alter cell
 PT proliferation
 XX
 XX Example 1; Page 62; 70pp; English.
 PS
 XX The present protein sequence is a mutant cyclin-dependent kinase
 CC inhibitor (CKI), p27. It is bound by hKIS, a serine/threonine kinase,
 CC that inhibits its ability to arrest cells in G1 phase. A substitution
 CC mutation (AGC to GCT) results in a serine to alanine (S10A) change in
 CC the protein, that abolishes phosphorylation of GST-p27, without affecting
 CC in vitro binding with hKIS. Endogenous p27 was detected in the cytoplasm
 CC and at higher levels in the nucleus. hKIS functions as an inhibitory
 CC kinase of CKI p27. The hKIS sequences are used to modulate cell
 CC proliferation and treat cell proliferative and vascular diseases.
 CC The polynucleotide sequence may be used in gene therapy to treat
 CC vascular disorders such as restenosis or atherosclerosis.
 XX
 SQ Sequence 198 AA;
 Query Match 82.7%; Score 91; DB 21; Length 198;
 Best Local Similarity 94.7%; Pred. No. 8.5e-06;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 2 NAGSVEWTPKKPGLRRQT 20
 ID AAY44400
 Db 180 NAGSVEQTPKKPGLRRQT 198
 RESULT 14
 AAY44400
 ID AAY44400 standard; Protein; 198 AA.
 XX
 AC AAY44400;
 XX
 XX 22-MAR-2000 (first entry)
 DT Human p27(Kip1) kinase inhibitor protein.
 XX
 DE p27(Kip1) kinase inhibitor protein; FKBP-12; p27(Kip1).FKBP-12 complex;
 KW cytosolic drug-binding protein; yeast two hybrid assay system;
 KW cell differentiation; apoptosis; neurodegeneration; tumorigenicity;
 KW cell proliferation related disorder; atherosclerosis; autoimmune disease;
 KW transplant rejection; inflammation; allergy; cancer; viral infection;
 KW membranous nephropathy; CDK; cyclin-dependent kinase.
 XX
 XX Homo sapiens.

```

XX FH Key Location/Qualifiers
XX FT Region 43..198
XX FT /note= "Prey sequence that interacts with FKBP-12"
XX
XX PN WO9965939-A1.
XX PD 23-DEC-1999.
XX PF 18-JUN-1999; 99WO-US13659.
XX XX 18-JUN-1998; 98US-0099857.
XX XX (CURA-) CURAGEN CORP.
XX XX
XX PI Nandabalan K, Yang M;
XX DR WPI; 2000-116763/10.
XX DR N-PSDB; AAZ29564.
XX
XX PT New complex of p27(Kip1) and FKBP-1. for treatment, prevention and
XX diagnosis of, e.g. cancer and autoimmune disease
XX
XX PS Claim 1; Fig 1; 78pp; English.
XX
XX CC The present sequence is p27(Kip1) kinase inhibitor protein. This protein
XX regulates CDK activity by inhibiting cyclin-CDK complex-associated kinase
XX activity. Interaction between p27(Kip1) and FKBP-12 (a cytosolic
XX drug-binding protein) to form a p27(Kip1).FKBP-12 complex was detected
XX using a modified yeast two hybrid assay system. This interaction is
XX used in regulating many cell functions, e.g. cell cycle progression,
XX differentiation, apoptosis, neurodegeneration, response to viral
XX infection, tumorigenicity. p27(Kip1).FKBP-12 complex and its
XX corresponding nucleic acid sequence is used in diagnosis and treatment
XX of cell proliferation related disorders. Specified diseases are
XX atherosclerosis; autoimmune diseases (e.g. transplant rejection,
XX inflammation or allergy); neurodegeneration; cancer; membranous
XX nephropathy and viral infections.
XX
XX SQ Sequence 198 AA;

Query Match 82.7%; Score 91; DB 21; Length 198;
Best Local Similarity 94.7%; Pred. No. 8.5e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 NAGSVEWTPKKPGLRRRQT 20
    ||||| ||||| |||||
Db 180 NAGSVEQTPKKPGLRRRQT 198

RESULT 15
AAB84649
ID AAB84649 standard; Protein; 198 AA.
XX
XX AC AAB84649;
XX
XX DT 05-SEP-2001 (first entry)
XX
XX DE Amino acid sequence of a human p27 KIP1 polypeptide.
XX
XX KW Knockout mouse; p19 INK4d; p27 KIP1; bradykinesia; cell growth;
XX KW proprioceptive abnormality; neuronal growth; motor disorder;
XX KW neuronal cell.
XX
XX OS Homo sapiens.
XX
XX PN US6245965-B1.
XX
XX PD 12-JUN-2001.
XX
XX PF 29-JAN-1999; 99US-0240906.
XX
XX PR 29-JAN-1999; 99US-0240906.

```

```

XX
XX PA (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX
XX PI Roussel MF, Smeyne R, Zindy F, Cunningham J;
XX
XX DR WPI; 2001-424332/45.
XX DR N-PSDB; AAH28328.
XX
XX PT New knockout mouse having a genome comprising a homozygous disruption
XX of both p19 INK4d and p27 KIP1 genes, useful as animal models studying
XX motor disorders having symptoms that include bradykinesia and
XX proprioceptive abnormalities
XX
XX PS Disclosure; Columns 31-34; 24pp; English.
XX
XX CC The present sequence represents a human p27 KIP1 polypeptide. The
XX specification describes a knockout mouse whose genome is manipulated
XX to comprise a homozygous disruption of both the p19 INK4d and p27 KIP1
XX genes, where homozygous disruption of these genes results in the knockout
XX mouse exhibiting bradykinesia and/or proprioceptive abnormalities. or
XX prevents in the expression of functional p19 INK4d and p27 KIP1 proteins.
XX The knockout mouse and cells may be used to identify potential
XX modulator of cell growth and more particularly neuronal growth. The
XX knockout mouse is useful as animal model for studying motor disorders
XX having symptoms that include bradykinesia and/or proprioceptive
XX abnormalities and/or seizures, and in identifying potential modulators
XX of motor functions. Cells from the knockout mouse may be used as a
XX potential source of differentiated neuronal cells, and for identifying
XX agonists and antagonists of neuronal cell growth.
XX
XX SQ Sequence 198 AA;

Query Match 82.7%; Score 91; DB 22; Length 198;
Best Local Similarity 94.7%; Pred. No. 8.5e-06;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 NAGSVEWTPKKPGLRRRQT 20
    ||||| ||||| |||||
Db 180 NAGSVEQTPKKPGLRRRQT 198

```

Search completed: August 27, 2003, 09:41:03
Job time : 475 secs